

FLOOD ESTIMATION FOR ROADS, BRIDGES AND DAMS

by

Mohamed Parak

Submitted in fulfilment of the requirements for the degree of
Master of Science in Engineering
in the
Civil Engineering Programme
University of KwaZulu-Natal
2007

Durban
June 2007

ABSTRACT

Flood estimation can be classified into two categories, i.e. flood prediction and flood forecasting. Flood prediction is used for the estimation of design floods, which are floods associated with a degree of risk of being equalled or exceeded. Predictions are needed for the design and construction of infrastructure that are at risk to flowing water. Flood forecasting is used for the estimation of flood flows from an impending and/or occurring rainfall event (i.e. the estimation of the magnitude of future flood flows with reference to a specific time in the future). These are needed by catchment and disaster managers for the mitigation of flood damage. The estimation of flood magnitudes for flood forecasting requires the specific knowledge of prevailing surface conditions which are associated with the processes of rainfall conversion into flood runoff. In order to best achieve this, a distributed model (in order to exploit remotely sensed data and capture the spatial scale of the phenomenon) is used to continuously update the surface conditions that are important in this conversion process.

This dissertation focuses on both flood estimation categories. In the first part of the dissertation, attention is given to the improvement of two simple event-based design flood prediction methods currently in use by design practitioners, namely the regional maximum flood (RMF) and the rational formula (RF) by comparison with statistically modelled historical flood data. The second part of the dissertation lays the theoretical and practical foundation for the implementation of a fully distributed physically-based rainfall-runoff model for real-time flood forecasting in South Africa. The TOPKAPI model was chosen for this purpose. This aspect of the research involved assimilating the literature on the model, testing the model and gathering and preparing of the input data required by the model for its eventual application in the Liebenbergsvlei catchment. The practical application of the model is left for a follow-up study.

DECLARATION

I, **Mohamed Parak**, declare that this dissertation is my own work carried out under the supervision of **Professor G.G.S. Pegram** in the Civil Engineering Programme, University of KwaZulu-Natal, Howard College, Durban. It has not been submitted in part, or in whole to any other university. Where use has been made of the work of others, it has been duly acknowledged in the text. This work is in accordance with the requirements of the University for the award of the MScEng degree in the branch of Civil Engineering.

Date: **06 June 2007**

Name: **Mohamed Parak**

Signature: MParak

As the candidate's supervisor, I have approved this dissertation for submission.

Date: **06 June 2007**

Name: **Geoffrey GS Pegram**

Signature: G.G.S. Pegram

ACKNOWLEDGEMENTS

I would like to thank the following people / institutes:

- *Professor Pegram*: for sharing his knowledge and guidance with regard to the related subject matter, for his support, understanding and patience through some difficult times, and for his encouragement and enthusiasm which proved to be enough for the both of us to see this research through to the end.
- *My Family*: for their love and support, for bearing my troubles with patience, and for never losing hope in my ability to finish.
- *Eastern Centre of Transport Development (ECOTD)*: for providing the financial assistance that made this research possible.
- *Jeffares and Green*: for allowing me the time to put the final document together.
- *"The Vlei Boys"*: for providing a thoroughly enjoyable work environment during my two years of full-time study.

TABLE OF CONTENTS

ABSTRACT	ii
DECLARATION	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	viii
LIST OF TABLES	xiii
PREFACE	xv
CHAPTER 1. INTRODUCTION	1
1.1. Approaches to design flood prediction	2
1.2. Approaches to flood-forecasting	3
1.3. Objectives of dissertation	5
1.4. Overview of chapters	6
1.5. Chapter summary	7
CHAPTER 2. DESIGN FLOOD PREDICTION	9
2.1. The regional maximum flood (RMF) method	9
2.2. The rational formula method	12
2.3. The runhydrograph method	16
2.3. Chapter summary	19
CHAPTER 3. A REVIEW OF THE REGIONAL MAXIMUM FLOOD (RMF)	20
3.1. Methodology and results	21
3.1.1. Flood record modelling	21
3.1.2. Return period of the RMF	22
3.1.3. Inclusion of landscape parameters in empirical formulas	25
3.2. Discussion of results	31
3.2.1. Return period of the RMF	31
3.2.2. Inclusion of landscape parameters in empirical formulas	32
3.3. Chapter summary	33

CHAPTER 4. THE RATIONAL FORMULA FROM THE RUNHYDROGRAPH ...	34
4.1. Methodology and results	34
4.1.1. Streamflow database	36
4.1.2. Rainfall database	36
4.1.3. Calibration of the runoff coefficients	38
4.1.4. Hydrograph time base-length	41
4.1.5. Validation of the calibrated runoff coefficients	42
4.2. Discussion of results	46
4.2.1. Calibration	46
4.2.2. Hydrograph time base-length	47
4.2.3. Validation	49
4.3. Chapter summary	49
CHAPTER 5. REAL-TIME FLOOD-FORECASTING USING RAINFALL-RUNOFF MODELS	51
5.1. Hydrologic modelling of catchments	51
5.1.1. Development of hydrologic models	52
5.1.2. Concepts of hydrologic models	53
5.2. Chapter summary	61
CHAPTER 6. TOPKAPI MODEL	63
6.1. Description of the model	64
6.1.1. Model assumptions	66
6.1.2. Soil water flow model	67
6.1.3. Overland and channel water models	75
6.1.4. Evapotranspiration model	82
6.1.5. Moisture accounting in each cell	85
6.1.6. Solution of the non-linear differential equations	90
6.1.7. Calibration	96
6.2. Chapter summary	97

CHAPTER 7. THE APPLICATION OF THE TOPKAPI MODEL	98
7.1. Data requirements	98
7.1.1. The Liebenbergsvlei catchment	100
7.1.2. Digital Elevation Model (DEM)	102
7.1.3. Soils map	117
7.1.4. Landuse map	120
7.1.5. Rainfall input	121
7.1.6. Data alignment	124
7.2. Test application	126
7.2.1. Input parameters	127
7.2.2. Soil store	128
7.2.3. Overland Store	131
7.2.4. Flow partitioning	133
7.2.5. Channel Store	135
7.2.6. Running the model	137
7.3. Chapter summary	149
CHAPTER 8. CONCLUSION	150
8.1. Summary and discussion of research	150
8.1.1. A review of the regional maximum flood (RMF)	150
8.1.2. The rational formula from the runhydrograph	151
8.1.3. TOPKAPI model	152
8.2. Chapter summary	153
REFERENCES	155
APPENDIX A	163
APPENDIX B	179
APPENDIX C	197
APPENDIX D	203

LIST OF FIGURES

- Figure 2-1.** Map of Southern Africa indicating the Maximum Flood Peak regions and their associated K-values (Kovacs, 1988). **p. 10**
- Figure 2-2.** A standard bivariate normal probability density function, with a cross correlation coefficient of 0.85, plotted with log-transformed observed flood peak-volume pairs in probability space (from Hiemstra and Francis, 1979: 14). The bold lines in the positive quadrant are the 10- and 100-year return period joint-exceedence contours. The dashed lines include a quadrant to the upper right, which on average will include 1% of the observations. **p. 17**
- Figure 2-3.** Joint flood peak and flood volume exceedence contours, in probability space for a peak-volume cross-correlation coefficient of 0.85 (from Hiemstra and Francis, 1979: 53). **p. 18**
- Figure 3-1.** Determination of the return period associated with the RMF for Region 5.2. The bold line is the RMF estimate (defined by the Francou-Rodier equation) and the thin lines (dotted, dashed and solid) are trend-lines fitted to the 50-, 100- and 200-year floods estimated from a probabilistic analysis of recorded annual peaks for catchments in Region 5.2. **p. 23**
- Figure 3.2.** Determination of the return period associated with the RMF for Region 5 (description as per Fig. 3-1). **p. 23**
- Figure 3-3.** Determination of the return period associated with the RMF for Region 4.6 (description as per Fig. 3-1). **p. 24**
- Figure 3-4.** A plot of the k -parameter of the GEV Distribution for each catchment against the regional K -value of the RMF (Kovačs, 1988) for Regions 4.6, 5 and 5.2. The mean GEV k -parameter together with its standard errors is also shown for each regional K . **p. 25**
- Figure 3-5.** The Klip River catchment (represented by gauge C1H002) and basin properties (derived from Petras & Du Plessis, 1987; Kovačs, 1988; Midgley *et al.*, 1990 and Parak, 2003). The definitions of the *mean river slope*, *Strahler basin order* and *Shreve magnitude* are also shown. Reference should be made to Pegram and Parak (2004) in Appendix A for the definitions of these and other parameters. **p. 27**
- Figure 3-6.** The 20-year flood (Q_{20}) vs. *catchment area* for the calibration of an empirical model. The number of catchments used here was 13. **p. 29**
- Figure 3-7.** The 20-year floods generated from the model calibrated in Fig. 3-6 vs. the probabilistically modelled floods for the same catchments. The number of catchments used in this validation was 12. **p. 30**
- Figure 4-1.** A comparison of the runoff coefficients c from Chow *et al.* (1988: 498) with those calibrated in this study $c_{(T)}$. The c -values plotted from Chow *et al.* are shown in thick bold lines and extend from the 2- to 500-year recurrence intervals. **p. 40**
-

-
- Figure 4-2.** A graph, for the purposes of validation, showing a plot in *log space* of the 10-year probabilistic rational formula flood peaks Q_{RF} versus the 10-year GEV modelled flood peaks Q_{GEV} . p. 44
- Figure 4-3.** A graph, for the purposes of validation, showing a plot in *log space* of the 50-year probabilistic rational formula flood peaks Q_{RF} versus the 50-year GEV modelled flood peaks Q_{GEV} . p. 45
- Figure 4-4.** A graph, for the purposes of validation, showing a plot in *log space* of the 200-year probabilistic rational formula flood peaks Q_{RF} versus the 200-year GEV modelled flood peaks Q_{GEV} . p. 45
- Figure 4-5.** The method employed by Hiemstra and Francis (1979) to extract independent hydrographs from a continuous flow record, showing that a lower truncation level is likely to provide a bigger volume. p. 48
- Figure 5-1.** Typical processes of the hydrologic cycle (<http://www.cet.nau.edu/Projects/SWRA/research.html>). p. 52
- Figure 5-2.** A radar image of rain depth (mm) accumulated for the 24-hour period from 08:00 21/06/2005 to 08:00 22/06/2005 (<http://metsys.weathersa.co.za/simar-archive.html>) superimposed on a digital elevation model (DEM) of Southern Africa (HYDRO1k, 1996). p. 57
- Figure 5-3.** A satellite image of rain depth (mm) accumulated for the 24-hour period from 08:00 21/06/2005 to 08:00 22/06/2005 taken from the Meteosat-8 Satellite (Pegram et al., 2005) superimposed on a digital elevation model (DEM) of Southern Africa (HYDRO1k, 1996). p. 57
- Figure 6-1.** Variation of suction pressure head ψ and hydraulic conductivity K with moisture content θ for an unsaturated clay soil (Chow et al., 1988:103). p. 68
- Figure 6-2.** The dimensions (in m) of the soil store in a pixel as defined by a DEM and as required by TOPKAPI. The picture is not to scale and the dimensions have been exaggerated. p. 73
- Figure 6-3.** The water balance for the soil store of a generic non "source" cell derived in Eq. 6-17 for the TOPKAPI model (from MUSIC Final Report, 2004: 53). p. 75
- Figure 6-4.** The rectangular cross section of an open channel flow of flow depth y and flow width B . p. 77
- Figure 6-5.** The water balance for the overland store of a generic cell derived in Eq. 6-21 for the TOPKAPI model (from MUSIC Final Report, 2004: 55). p. 79
- Figure 6-6.** The water balance for the channel store of a "source" cell derived in Eq. 6-23 for the TOPKAPI model. p. 81
- Figure 6-7.** The tree-shaped network form by the cells of the TOKAPI model (from MUSIC Final Report, 2004: 53). p. 86
- Figure 6-8.** A generic catchment consisting of three cells showing the intra- and inter-cell operations of the TOPKAPI model. p. 88
-

- Figure 7-1.** The Liebenbergsvlei catchment in South Africa showing the eight quaternary sub-catchments and the river network, which is shown at a spatial detail of 1:250 000 (Midgley et al., 1994). Also shown are the locations of the 45 telemetering raingauges, 9 flowgauges and the S-Band MRL5 weather radar. **p.101**
- Figure 7-2.** A DEM of Africa, South Africa and the Liebenbergsvlei catchment at a grid resolution of 1km square from the geographic database of the United States Geological Survey (HYDRO1k, 1996). **p.102**
- Figure 7-3.** The *D8 flow model* of O'Callaghan and Mark (1984) as applied in ARCGIS™. **p.104**
- Figure 7-4.** A flow direction raster showing the direction code of each cell determined using ARCGIS™ for a 200m square resolution DEM of the Liebenbergsvlei catchment. **p.104**
- Figure 7-5.** A flow accumulation raster showing the number of upslope cells flowing into each cell determined using ARCGIS™ for the Liebenbergsvlei catchment. **p.105**
- Figure 7-6.** A comparison between a stream network delineated for the Liebenbergsvlei catchment from a DEM, using a threshold value of 500 cells and a stream network digitised from a topographic map (at a spatial scale of 1:250 000 from Midgley et al. (1990)) . The DEM delineated network shown on the left, although appearing disjointed in the image, is continuous in reality. **p.106**
- Figure 7-7.** A 218m square pixel resolution DEM of the Liebenbergsvlei catchment is shown on the right (from DLSI, 1996). The pixels of this DEM were "resampled", using an inbuilt function on ARCGIS™, to a resolution of 1000m square (shown on the left). **p.108**
- Figure 7-8.** The D8 flow direction raster codes resolved into a D4 flow direction raster. **p.110**
- Figure 7-9.** On the left is shown a D4 flow direction output as a result of resolving a D8 direction raster of the Liebenbergsvlei catchment into the four cardinal directions only. The legend of the direction raster show the four cardinal direction codes as well as an error code (-100 000) to mark those cells whose resolved directions flowed out of the catchment. On the right is shown the resulting flow accumulation raster based on the D4 method. **p.111**
- Figure 7-10.** Anomalous positions for direction codes that resulted from the resolving of a D8 flow direction raster into a D4 raster. These pixels created areas of internal drainage within the catchment. **p.111**
- Figure 7-11.** The "corrected" flow accumulation raster is shown on the left which is overlain on the right with a stream network (in red) traced from a topographic map (from Midgley et al., 1994) at a spatial scale of 1:250 000. **p.112**
- Figure 7-12.** The stream networks delineated utilizing five different threshold areas, i.e. 5, 10, 25, 50 and 65km square, on ARCGIS™. The superimposition of the 1:250 000 network (from Midgley et al., 1994) on the 65km square network is shown on the bottom right. **p.114**

- Figure 7-13.** A surface slope raster (in degrees) is shown computed from the 1km DEM on ARCGIS™. The resolution of the raster is 1km square. **p.116**
- Figure 7-14.** The stream orders, ordered according to the method of Strahler (1964), based on a stream network extracted using a threshold area of 25km² on ARCGIS™. **p.117**
- Figure 7-15.** A vector-based soils map from SIRI (1987) was converted into a raster with a pixel resolution of 1km square from which the raster-based soils map of the Liebenbergsvlei catchment was clipped. The legend distinguishes soils groups of which the properties are given in Table 7-2. **p.118**
- Figure 7-16.** A landuse map of Africa, South Africa and the Liebenbergsvlei catchment, all at a pixel resolution of 1km square (GLCC, 1997). **p.121**
- Figure 7-17.** An instantaneous volume scan of radar reflectivity (dBz) at 2km altitude above ground level (a.g.l.), from which the rainfall for the Liebenbergsvlei catchment has been clipped out. The resolution of the images is 1km square. **p.123**
- Figure 7-18.** A screen capture of the geographic and projected coordinate system details, from ARCGIS™, which has been used uniformly in this research to represent all data covering the extent of the Liebenbergsvlei catchment. **p.125**
- Figure 7-19.** A screen capture of the headings of the files of all the input data, in text-format, which describe the properties of each pixel of the Liebenbergsvlei catchment, which is required by the TOPKAPI model. These data can be displayed in ARCGIS™ in raster-format. The figure shows that all the data consist of 62 columns, 121 rows, are aligned at the lower left corner in a projected coordinate system (see Fig. 7-18) and are at a cell size resolution of 1000m square. **p.126**
- Figure 7-20.** The operations of the soil store of Cell 1 for the first ten time steps as modelled using Microsoft Excel™. **p.129**
- Figure 7-21.** The operations of the overland store for Cell 1 as modelled on Microsoft Excel™. **p.132**
- Figure 7-22.** The flow partitioning operations of Cell 1, where flow is partitioned between the channel store of Cell 1 and the soil store of the downstream cell, i.e. Cell 2. **p.135**
- Figure 7-23.** The operations of the channel store of Cell 2 as modelled on Microsoft Excel™. **p.136**
- Figure 7-24a, b and c.** The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 1. **p.145**
- Figure 7-25a, b and c.** The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 2. **p.146**
- Figure 7-26a, b and c.** The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 3. **p.147**
- Figure 7-27a, b and c.** The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 4. **p.148**

Figure D-1. The soil store model for Cell 1.	p.204
Figure D-2. The overland store model for Cell 1.	p.205
Figure D-3. The flow partitioning model for Cell 1.	p.206
Figure D-4. The channel store model for Cell 1.	p.207
Figure D-5. The soil store model for Cell 2.	p.208
Figure D-6. The overland store model for Cell 2.	p.209
Figure D-7. The flow partitioning model for Cell 2.	p.210
Figure D-8. The channel store model for Cell 2.	p.211
Figure D-9. The soil store model for Cell 3.	p.212
Figure D-10. The overland store model for Cell 3.	p.213
Figure D-11. The flow partitioning model for Cell 3.	p.214
Figure D-12. The channel store model for Cell 3.	p.215
Figure D-13. The soil store model for Cell 4.	p.216
Figure D-14. The overland store model for Cell 4.	p.217
Figure D-15. The flow partitioning model for Cell 4.	p.218
Figure D-16. The channel store model for Cell 4.	p.219

LIST OF TABLES

- Table 3-1.** Summary of the trend-lines from Figs. 3-1, 3-2 and 3-3. The coefficients of determination (R^2 -values) appears in parentheses. **p. 24**
- Table 3-2:** Results of the *step-wise* regression of the model calibration and validation. The numbers in parentheses flag the "best" fit (based on the R^2 -value) to the validation data. **p. 30**
- Table 4-1.** The results of the calibration of the c -coefficient of the rational formula on flood peak and flood volume pairs from Hiemstra and Francis (1979). **p. 39**
- Table 4-2.** Runoff coefficients for use in the rational method for undeveloped (rural) regions in Austin, Texas in the USA (from Chow et al., 1988: 498). **p. 40**
- Table 4-3.** The mean and standard deviations of the ratio of the hydrograph time base-length B to the catchments' time of concentration T_c as a function of recurrence interval T . The proportion of B/T_c values above 3 in each interval is given in the third row. **p. 42**
- Table 4-4.** Equivalent land coverage types from the descriptions of Petras and du Plessis (1987) and Chow et al. (1988: 498) for the catchments used in validation. **p. 43**
- Table 4-5.** A summary of the *power-law* curves, of the form $Q_{RF} = aQ_{GEV}^b$, fitted to the graphs of Q_{RF} vs. Q_{GEV} (where Q_{RF} are the flood peaks obtained from the probabilistically applied rational formula and Q_{GEV} are the statistically modelled flood peaks). The *coefficient of determination* (R^2 -value) and the average ratio of Q_{RF}/Q_{GEV} for each recurrence interval are given in the last two rows respectively. **p. 44**
- Table 5-1.** General data requirements for catchment models (from Singh and Woolhiser, 2002). **p. 55**
- Table 7-1.** The drainage densities of the extracted stream networks shown in Fig. 7-12. **p.113**
- Table 7-2.** The related soil properties of the landtype map (SIRI, 1987) of the Liebenbergsvlei catchment (see Fig. 7-15). **p.119**
- Table 7-3.** Input parameters chosen for the "four cell generic catchment" for the test application of the TOPKAPI model on Microsoft Excel™. The suggested range of the parameters (in Liu and Todini, 2002) is given in parentheses in the second column. **p.128**
- Table 7-4.** The range of values that λ can take when regulating the flow Q to the channel of a cell and to the downstream cell. **p.134**
-

Table 7-5. A summary of the “continuity check” on each store of each cell for the “four cell generic catchment” modelled on Microsoft Excel™. The values in the last column were computed from Eq. 7-1. **p.140**

Table 7-6. A summary of the “corrected continuity check” on each store of each cell for the “four cell generic catchment” modelled on Microsoft Excel™. A “lack of fit” factor was calibrated for each store of each cell (last column) to force continuity to hold. **p.142**

PREFACE

The research presented in this dissertation is divided into two parts, i.e. design flood prediction in the first part and real-time flood-forecasting in the second part. Although both parts of this research generally fall under the banner of *flood hydrology*, they are conceptually different. The reason for choosing essentially different topics for this research was based primarily on the source of funding, which was provided by the Eastern Centre of Transport Development.

The first part of this research focuses on flood prediction for the estimation of design floods. Design floods are floods associated with a degree of risk of being equalled or exceeded. It was the intention of this study to provide a useful guide or modification of established flood prediction methods that could be directly used by practitioners charged with the design of structures that are at risk to flowing water. Structures such as dams, river bridges, roads and floodplain developments are a few types of the structures that fall into this category. It is evident that this aspect of the research has a direct influence on the transport sector, where design floods are used for road and bridge design. The research conducted in this first part resulted in two publications of which Chapters 2, 3 and 4 are the focus.

It was found that the first part of this research was not enough for a stand-alone masters dissertation and at the same time there was a pressing need to examine the efficacy of a candidate distributed rainfall-runoff model for flood-forecasting purposes. Flood-forecasting is the estimation of the magnitude of future flood flows with reference to a specific time in the future and is used for the mitigation of damage caused by floods. Thus it was decided in this research to also focus on this aspect of flood hydrology, and this forms the focal point of the second part of this dissertation (Chapters 5, 6 and 7). The first part is essentially "looking back", through the review of established flood prediction methods. The second part is essentially "looking forward" to hydrology in the 21st century in order to investigate the use of advanced computational capabilities to seamlessly integrate the various inputs from remote sensing techniques and spatial data towards distributed rainfall-runoff modelling. Although there is no direct contribution to the transport sector from the second part of this research, the mitigation of flood damage (social and economic) through the use of flood-forecasting techniques does have an indirect influence on this sector.

CHAPTER 1

1. INTRODUCTION

Flood estimation can be broadly classified into two categories, i.e. design flood prediction and flood-forecasting. Design flood prediction, as understood in this research, is concerned with the estimation of the magnitude of flood events and the probability of those flood events being equalled or exceeded. These estimates are needed for the planning and design of engineering projects that are at risk to flood water. Catastrophic floods have a huge economic, social and environmental impact and thus reliable design flood prediction is a subject of great importance. However, consistent estimates of design floods remain a current challenge in hydrology (Smithers and Schulze, 2001).

Flood-forecasting, as opposed to prediction, means the estimation of flow conditions at a specific time in the future; prediction is the estimation of future conditions without a reference to a specific time (Lettenmaier and Wood, 1992). Flood-forecasting is thus the estimation of expected future flood flows and the precision or uncertainty associated with the forecast. Forecasts are needed for improved warnings and operational decisions for the mitigation of flood damage. More formally, flood-forecasting involves the provision of reliable, intelligible forecasts of flood flows with long *forecast lead-times* (which is catchment dependent but generally greater than 12 hours) and explicit error bounds, made available at frequent intervals to hydrological operators, decision makers and disaster managers (Pegram, 2003b).

Flood forecasting is an essential tool for catchment and disaster managers for the provision of accurate and reliable forecasts of future flood flows for the mitigation of flood damage. In order to achieve this, the forecasts need to be delivered with a *sufficient lead-time* so that any *mitigation operations may be implemented*. To maximise lead-time, precipitation information is needed in real-time, or forecasted ahead of time, to take advantage of the delay it takes the precipitation to reach the point of interest on the stream or channel from where it falls on the ground. To this end, a rainfall-runoff model is needed to simulate the process that occurs in converting precipitation into flood runoff.

1.1 Approaches to design flood prediction

There are three primary approaches to design flood prediction: *empirical, deterministic and probabilistic*. The approach that is used depends largely on the historical data (namely precipitation and streamflow records) that are available at the site. To a lesser extent site-specific considerations, such as catchment size and the nature of the design project, also dictate the method that should be used. However, owing to the uncertainty in flood predictions, no method should be used in isolation and all three approaches should be used where possible (Görgens, 2002).

An empirical approach attempts to predict the flood peak based on catchment and regional characteristics only. This method is used if the catchment is ungauged and there exists no historic precipitation and streamflow data. Thus, this method uses regional characteristics and some descriptor of catchment morphometry (such as catchment area) to predict flood magnitudes. Based on previous calibrations at other sites, a probability of exceedence (or return period) can usually be associated with this estimate. One such approach used in South Africa is the regional maximum flood (RMF) method developed by Kovačs (1988).

Deterministic methods attempt to replicate all the factors involved in flood production, i.e. in the conversion of rainfall into flood runoff. These methods require historic precipitation in order to estimate the design storm associated with a given exceedence probability. Through the simplification of the rainfall-runoff process, deterministic methods then convert the design storm into the design flood of supposedly the same exceedence probability (ignoring the effects of joint probability). This approach can also be used on ungauged catchments (where no precipitation and streamflow records exist) if suitable regional techniques can be used to predict the design storm for that location. Methods such as the rational formula (Mulaney, 1850) and the SCS method (Soil Conservation Service, 1972) are examples of commonly used deterministic methods for small to medium sized catchments. Both these methods have since been updated to suit local conditions, for the rational formula by Alexander (2002) and Pegram (2003a), and for the SCS method by Schmidt and Schulze (1987a and 1987b) and Schmidt et al. (1987).

The probabilistic (or statistical) approach bases its flood peak estimate for a given catchment on the fitting of the most appropriate probability distribution to flood records from the catchment. The reasoning behind the concept of a statistical approach is that

floods can be viewed as a random process and as such can be described by a probability distribution. Distributions such as the Log-Normal (LN) Distribution, the Log-Pearson-III (LPIII) Distribution and the General Extreme Value (GEV) Distribution are commonly used to model flood frequencies. The runhydrograph method of Hiemstra and Francis (1979) was one such attempt to calibrate a statistical method for design flood estimation in South Africa. A probabilistic approach, theoretically, is the most consistent approach to estimate design floods provided that the site of interest be at or near a flow gauge (which is seldom the case) and a long and reliable flow record exists from that gauge (Görgens, 2002).

In the estimation of design floods, both the empirical and the deterministic approaches endeavour to behave in a probabilistic manner, i.e. to predict a flood magnitude with an associated probability of exceedence. This point is not clearly obvious, since the traditional application of a deterministic method, such as the rational formula for example, attempts to reproduce a historic event, i.e. a flood flow from a given set of rainfall and catchment conditions (storm patterns, ground cover conditions, antecedent moisture conditions, etc.). In a probabilistic sense, all the variables of the rational formula would need to be associated with a probability of exceedence in order to derive a flood flow of the same exceedence probability (again ignoring the effects of joint probability). Thus, a probabilistic approach to the empirical and deterministic methods is used to estimate the magnitude of the peak discharge from a site for a given probability of exceedence. This peak should be equivalent to a discharge estimated from a frequency analysis of flood records if a long and representative record were available at that site.

1.2. Approaches to flood-forecasting

Forecasting of flood flows usually falls within the realm of *short-term forecasts*, where the forecast lead-times are less than seven days (Lettenmaier and Wood, 1992). The forecast lead-time is the time interval for which the forecast is made and ranges from a few hours to a few days, depending on the type of forecast model used and the size of the catchment. As the lead-time increases and the area reduces, the accuracy of the forecast usually decreases, where the accuracy refers to the difference between the amount forecast and the actual amount that occurs. *Long-term forecasts* are typically meteorologically based and are issued usually by weather services with lead-times of

up to several months. They are more often used for the management of water resources and are based on seasonal behaviour and global trends of climate change. There are various components of the flood-generation process, within the realm of short-term flood-forecasting, for which forecasts can be made. The forecasts for each component are associated with different levels of forecast accuracy and forecast lead-time. In this study these components were categorized according to the discipline into which they fall, namely *meteorologic*, *hydrologic* and *hydraulic*, which coarsely describe the water in the atmosphere, over the land and in the channel respectively. The approach that is usually used depends on the intended application of the forecast information.

Meteorologic forecasts are (from the hydrologic view point) concerned with the prediction of precipitation information ahead of its time. This is achieved through the use of quantitative precipitation forecasts (QPF) based on numerical weather prediction systems (Bartholmes and Todini, 2003). The forecast precipitation information is coupled with a rainfall-runoff model in order to transform the rainfall into runoff. In this instance, the greatest forecast lead-time is achieved and can be as much as four days in advance, but the accuracy of the forecast is diminished. This type of forecast is usually referred to as *medium-range flood-forecasting* on account of the lead-time achieved.

Hydrologic forecasts are concerned with the simulation of the rainfall-runoff process in real time, i.e. as it happens. These forecasts use real-time observations of precipitation from rain gauges, weather radar and satellite observations (or a combination of these) coupled with a rainfall-runoff model. The forecast lead-time is extended to the time it takes the precipitation to cover the distance between the area where it falls and the problem site on the stream channel (Bartholmes and Todini, 2003). The forecast lead-time is thus dependent on the size of the catchment and its response time; the latter variable is a function of the catchment's physical characteristics such as slope, soil, geology and landcover. Real-time flood-forecasting is associated with greater forecast accuracy than medium-range flood-forecasting.

Hydraulic forecasts involve the inference of river level (or discharge) at a downstream section/station on the basis of stage (or discharge) at a point upstream (Reed, 1984). This is achieved through flood routing and provides the most precise forecast available with the lowest associated forecast uncertainty. However, forecast lead-times are limited to the travel times of the flood-wave in a channel (Bartholmes and Todini, 2003).

1.3. Objectives of dissertation

In the first part of this dissertation (Chapters 2, 3 and 4), two simple event-based flood prediction methods, which are currently used by design practitioners, are investigated. These are the regional maximum flood (RMF) and rational formula methods. The goal of this aspect of the research is to assist designers in using these methods by investigating issues related to the application of these methods.

The RMF (Kovács, 1988) is an empirically derived upper limit flood peak that can reasonably be expected at a given site. It predicts the regional “maximum” flood that can be expected from a given site based only on the site’s catchment area and location. The advantages of this empirical method is its ease of use as it deals directly with the variable of interest, namely the flood peak discharge, and its avoidance of the assumptions involved in transforming rainfall inputs into flood outputs.

A disadvantage of the RMF is that a return period cannot reliably be associated with its peak estimate. Although Kovaács (1988: 19) estimated the return period to be greater than 200-years, others such as Görgens (2002) suggest that the return period of the RMF is actually much larger than 200-years. Thus the return period of the RMF was assessed in this study as part of the research into flood-prediction methods. This was accomplished by comparing flood magnitudes determined from the RMF method with statistically modelled floods, of known return period, for the same catchments. Furthermore, the dependence of empirical methods (such as the RMF) on catchment area as the main independent parameter of flood computation was also investigated. This was accomplished by calibrating empirical formulae on other morphometric variables of the landscape including catchment area and comparing these models with formulae calibrated on catchment area only.

The rational formula is possibly the most widely used method for predicting design floods from design storms for urban catchments and small (<15km²) rural catchments, despite its criticism regarding its over-simplification of complex hydrological processes. One of the main criticisms of the rational formula regards the difficulty faced by practitioners in the probabilistic estimation of the runoff coefficient of the formula. With this in mind, a calibration of this variable, on statistically derived design flood peak and volume pairs for catchments in South Africa, was investigated in this research. The runhydrograph method of Hiemstra and Francis (1979) was used to derive the design flood peak and volume pairs for the calibration. The significance of this calibration to

designers is tested by attempting to extend the calibrated variables to ungaged catchments. This latter aspect is also investigated in this research in an effort to validate the calibrated coefficients.

In the second part of this dissertation (Chapters 5, 6 and 7) the focus is shifted to flood forecasting, namely in the assessment of the efficacy of a candidate rainfall-runoff model for the simulation of catchment hydrology. The simulation of catchment hydrology can be performed at various time and spatial scales. However, in order to exploit the type and quality of data currently available, models for real-time flood-forecasting need to be physically based and distributed. Thus the chosen model for this aspect of the research is one such model and is the TOPKAPI model of Liu and Todini (2002).

Since this model is novel to South Africa, this aspect of the research firstly involved the assimilation of disparate sources of literature on the model and the systematic combination of this information into a coherent whole for presentation as a candidate rainfall-runoff model for flood-forecasting purposes. After dissecting and resolving issues that were not clearly obvious from the literature, the model was then tested in a generic environment and the input data required by the model for its eventual application in the Liebenbergsvlei catchment was gathered and prepared. The practical application of the model is left for a follow-up study. The goal of this part of the research is to lay the theoretical and practical foundation for the implementation of a fully distributed physically-based rainfall-runoff model for real-time flood-forecasting in South Africa.

1.4. Overview of chapters

- *Chapter 2:* Introduces the primary concepts of design flood prediction and the theory of three prediction methods, two of which are reviewed in Chapters 3 and 4 respectively.
- *Chapter 3:* Provides a review of the empirically-based regional maximum flood (RMF) method (Kovačs, 1988) by estimating a return period associated with its peak discharge and by examining the effect of the inclusion of additional geomorphological parameters in the area-based empirical equation.

- *Chapter 4:* Provides a review of the rational formula method of flood prediction by calibrating the runoff coefficient of the formula on characteristic past flood peak and volume pairs of the runhydrograph method (Hiemstra and Francis, 1979).
- *Chapter 5:* Introduces the theory and concepts associated with flood-forecasting through the use of distributed rainfall-runoff models in simulating catchment hydrology.
- *Chapter 6:* Provides a comprehensive description of the candidate distributed rainfall-runoff model chosen for this study, namely the TOPKAPI model, by synthesizing and dissecting the literature on the model.
- *Chapter 7:* Describes the preparatory work that was performed in gathering and manipulating the input data (for the Liebenbergsvlei catchment) required by the model as well as the test application of the model in generic circumstances to confirm the model's operations.
- *Chapter 8:* Provides a review of the investigations carried out in this research, an appraisal of the value added by it and the points of departure for follow-up studies.

1.5. Chapter summary

It is hoped that that the research presented in this dissertation has made the following contributions to the engineering and hydrologic community. Firstly, with regard to the review of simple event-based flood prediction methods, it was discovered that the magnitude of the regional maximum flood (RMF) closely approximates a statistically derived 200-year flood. Thus in a design situation, it was felt that to ascribe a 200-year return period to the RMF would be reasonable. It was also found that the use of *catchment area*, as the sole landscape parameter of an empirical model, provided the best predictions of floods (based on the data used) when compared to empirical models utilising other measures of the landscape as independent variables. Thus the use of *catchment area* as the sole independent variable in calibrating empirical equations proved to be most practical (as it was easiest to quantify) and efficient for the purposes of flood estimation. Secondly, with regard to the calibration of the rational

formula's runoff coefficient (c), it was discovered that tabulated values of c from Chow et al. (1988: 498) proved reasonable for use in design checks of large catchments as well as small.

In the second part of this dissertation, the contribution made by this study was firstly in the introduction, analysis and explanation of a chosen rainfall-runoff model for flood forecasting purposes, i.e. the TOPKAPI model. This involved an intensive dissection of disparate sources of literature on the model and the systematic combination of this information as a coherent whole. Secondly, the input data required for the model was gathered and prepared through a geographical information system (GIS) and a test application was performed to verify the models operations. The research carried out here identified and resolved certain issues with regard to the models practical application. Issues such as the generation of TOPKAPI specific input were explained, the setup of the model's operations and the running of the model were performed. Since this model is novel to South Africa, the contribution made by this aspect of the research is the essential first step in laying the foundation for the models actual implementation, which is to be achieved in a follow up study.

CHAPTER 2

2. DESIGN FLOOD PREDICTION

This chapter is intended to briefly introduce the theory of three design flood prediction methods, namely the regional maximum flood (RMF), the rational formula and the runhydrograph method. The shortcomings of the former two methods are also discussed in preparation for the review of these methods in Chapters 3 and 4 respectively. These reviews formed the basis of two journal articles towards this research, namely Pegram and Parak (2004) and Parak and Pegram (2006). These articles are attached in Appendices A and B respectively.

2.1. The regional maximum flood (RMF) method

The RMF (Kovačs, 1988) is based on the use of the Francou-Rodier (Francou and Rodier, 1967) equation (Eq. 2-1) to define flood peak envelope curves. Kovačs (1988) used this equation, together with 519 observed flood peaks from catchments in Southern Africa, to delimit hydrologically homogeneous regions. When these peaks were plotted against catchment area for each of the regions, it was discovered for catchments larger than 100km² that the plots defined an upper limit of expected flood peaks for a given region. The Francou-Rodier equation is given as:

$$Q = 10^6 \left(\frac{A}{10^8} \right)^{1-0.1K} \quad (2-1)$$

where Q is the flood peak in m³/s, A is the catchment area in km² and K is a regional factor (dimensionless) which is indicative of the flood magnitude potential of the area, shown in Fig. 2-1. K accounts for the influences of variations in rainfall (intensity, area and duration) and catchment characteristics (such as geology, land-form, vegetation cover, etc) in flood production. The RMF for a particular site is then computed from Eq. 2-1 based on the knowledge of the size of the catchment (area) and its location (to determine its regional K -value).

It should be noted at this juncture, that the “secret” to the success of the RMF is the careful way in which Kovačs chose the regions to group the flood data. He did this by examining the actual K -value (determined from Eq. 2-1) for each of the flood peaks used in his study. Regional boundaries of K were delimited by considerations of individual K -values within the region, the number and accuracy of the data in a particular area, existing boundaries, maximum recorded 3 day storm rainfall, topography, catchment orientation with respect to dominant storm generating weather systems, general soil permeability, main drainage network and the location of large dams situated upstream from the gauging sites (Kovačs, 1988: 9). Of these considerations, individual K -values were evidently the most important and the regions were traced based on this. In areas of high flood magnitude potential, a difference in K of 0.2 between adjacent regions was allowed for and a difference of 0.6 in areas of low flood magnitude potential.

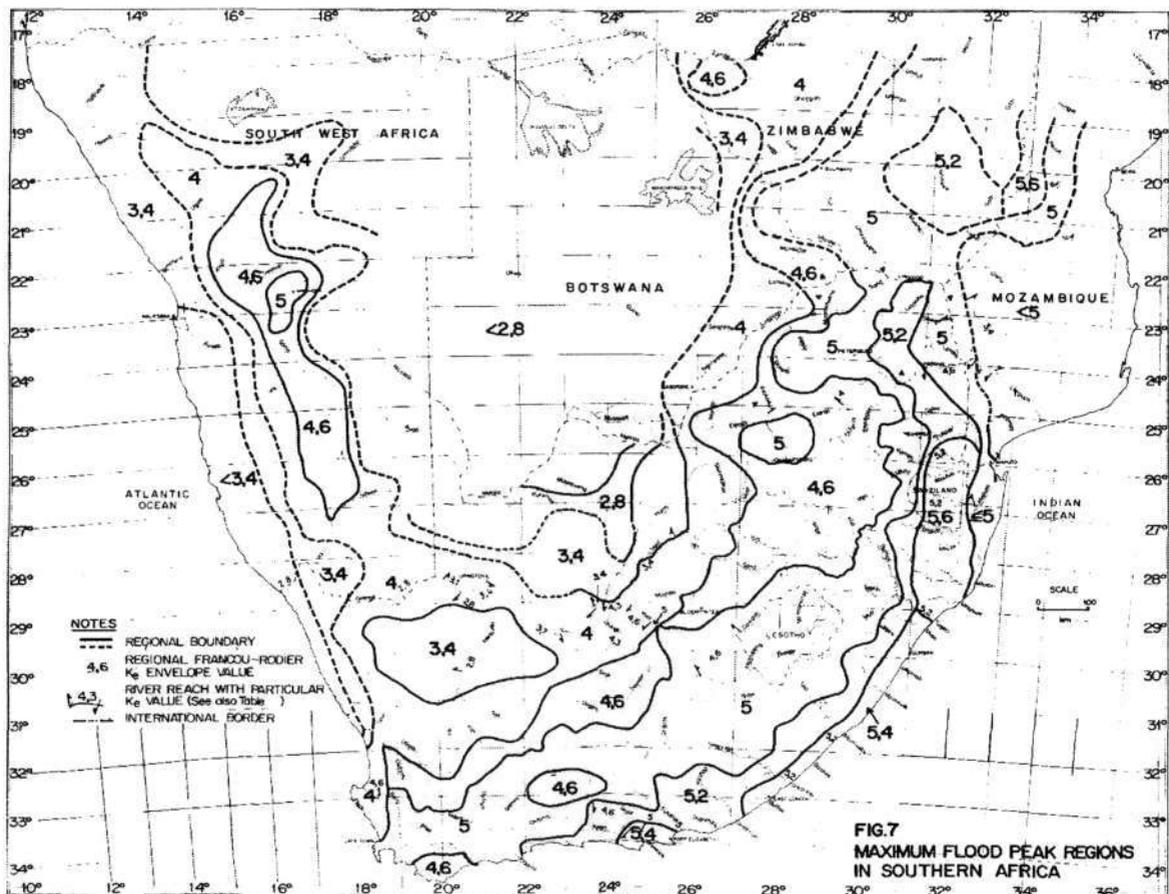


Figure 2-1. Map of Southern Africa indicating the Maximum Flood Peak regions and their associated K -values (Kovačs, 1988).

A disadvantage of the RMF method is that it does not clearly embody a design flood, i.e. a return period cannot easily be associated with its peak estimate. Kovačs himself

estimates the return period to be greater than 200 years (Kovačs, 1988: 19). He attempted to rectify this situation by modelling all the observed flood peaks of a particular *K*-region as a statistical sample. He used the sum of all the individual representative periods *N* (length of records at the gauges) of the peaks in that region as being equivalent to the overall recurrence interval of the particular *K*-region (station-year concept). Where the representative period of a flood was not known, Kovačs did not allow this to exceed 200 years and a provisional *N*-value was estimated based on the assumption that the ratio of the 200-year peak to RMF, Q_{200}/RMF , was 0.65. He then estimated ratios of the 50-, 100- and 200-year recurrence interval floods to the RMF for each region.

However, Görgens (2002) states that, "statistically speaking" Kovačs' (1988) method of determining the recurrence intervals was too simplistic. In reviewing Kovačs' method, Görgens recommends that the 50-, 100- and 200-year ratios "may need to be factored down by 0.7, 0.8 and 0.9 respectively". This implies that the RMF peaks have return periods that are actually much larger than 200-years, as opposed to the original estimation of Kovačs. As such Görgens suggests that the RMF method of Kovačs should be seen as the conservative or upper limit flood estimate in each recurrence interval.

Thus, the estimation of a return period which can reliably be associated with the RMF's peak estimate was, *inter alia*, investigated in this research. In order to accomplish this, floods determined from the RMF method were coaxially plotted with historical floods modelled with a statistical distribution (the GEV Distribution in this instance) from the same catchment. The flood records used were annual peaks from 130 catchments across South Africa and were actually a subset of the data set used by Kovačs (1988) in his study. The lengths of record of the data set used in this study ranged from 9 to 76 years. This investigation is explained in Chapter 3 and formed part of a journal article (Pegram and Parak, 2004), included in Appendix A.

The RMF method, like many other empirical methods, relies on *catchment area* as the independent variable in its flood estimate. However, one might expect other parameters of the fluvial landscape to play just as important a role in flood estimates and hence flood response. Flood geomorphologists, such as Horton (1932; 1945) and Strahler (1952; 1964) and many others since have been interested in relating flood discharges to physical measures of the landscape (morphometry). They identified parameters of the fluvial landscape which intuitively would correlate well with flood

discharge. Linear parameters (such as stream orders and stream lengths), areal parameters (such as catchment area, catchment shape and drainage density) and relief parameters (such as catchment relief, catchment slope, channel slope and ruggedness number) are some of the physical measures that have been identified as significantly affecting flood response. Instinctively, one can expect a multi-variable relationship between flood discharge and catchment morphometry to exist because a catchment is effectively “an open system trying to achieve a state of equilibrium” (Strahler, 1964). Precipitation and energy are inputs to the system and soil (eroded material) and excess precipitation leave the system through the catchment outlet. Within this system an energy transformation takes place converting potential energy of elevation into kinetic energy where erosion and transportation processes result in the formation of topographic characteristics. Thus it is evident that floods, and the landscape through which they drain, form a mutual relationship and ultimately catchment morphometry should reflect the long-term effects of this energy conversion phenomenon.

Thus, in addition to the estimation of a return period that could be associated with RMF-based estimates, research towards determining if landscape parameters, together with *catchment area*, improved the prediction of flood peaks in empirically based models was also conducted. This part of the investigation also formed part of the journal article Pegram and Parak (2004) which is included in Appendix A. This investigation is explained in Chapter 3.

2.2. The rational formula method

The rational formula was first proposed by the Irish engineer Mulvaney (1850) and has possibly become the best known and most widely used method for the determination of peak flood flows from rainfall events. It has survived numerous criticisms regarding its over-simplification of the complex hydrological processes of flood production but nonetheless is perhaps the most favoured method used by practitioners for peak flood estimation. The rational formula owes its popularity to the fact that it is easy to understand and simple to use. The peak flood flow due to a rainfall event on a catchment, determined from the rational formula, is expressed (in SI units) as:

$$Q_{RF} = ciA/3.6 \quad (2-2)$$

where Q_{RF} is the flood peak in m^3/s , c is the runoff coefficient, which (in the traditional deterministic approach) is defined as the proportion of precipitation that contributes to runoff, i is the storm rainfall intensity in mm/hr and A is the catchment area in km^2 .

The criticisms concerning the rational formula, in the form shown above, are not unfounded and the use of this method carries valid cautions that are based on the following assumptions built into the formula (which are not always explicit in its presentation):

- The maximum rate of runoff from a catchment is achieved when the duration of rainfall is equal to the time of concentration (T_c) of the catchment, which is defined as the time taken for the outflow from a catchment to reach near equilibrium due to rainfall uniformly spread in space and time.
- The spatial and temporal characteristics of rainfall are consequently ignored and the storm rainfall, as input into the formula, is assumed to be a rectangular pulse of duration T_c , deposited in lumped form on the catchment (i.e. there is no routing component implicit in the formula).
- The effects of joint probability are ignored and it is assumed that a T -year recurrence interval storm will produce a flood of the same recurrence interval, if the catchment is at 'average' conditions. This assumption is shared by most deterministic methods. However, it was shown by Gray (1973), in comparing the recurrence intervals of large historical storms and their resulting flood peaks, that the means of the two sets of recurrence intervals were closely matched, thus removing the necessity for the account of joint probability in deterministic methods.

The rational formula was previously limited in its application in South Africa to small catchments less than $15km^2$ in size (HRU, 1972) and it was only to be used as a check method, i.e. it was not to be used in isolation. It was also noted that sound engineering experience and judgment was required for its use. However, work that has since been done, locally by Alexander (2002) and Pegram (2003a), and abroad in Australia (Institute of Engineers Australia, 1987), has shown that these cautions were too conservative and its use may well be extended beyond small catchments.

As stated in Section 1.1, a probabilistic approach to the rational formula is needed for the estimation of design floods. In this case, the variables c and i (the runoff coefficient and rainfall intensity respectively) of the formula need to be associated with a probability of exceedence. Pilgrim and Cordery (1992) have stated that the design

situation is exactly suited to the probabilistic approach of the rational formula and has little similarity with the deterministic rational formula, and so the criticisms associated with the deterministic approach are not necessarily valid for the probabilistic design case. Alexander (1990) stated that as the catchment size increases the value of c becomes more probabilistic than deterministic in its derivation. The probabilistic approach to the rational formula has the same form as Eq. 2-2 but is defined more specifically as:

$$Q_{(T)} = c_{(T)} i_{(T_c, T)} A / 3.6 \quad (2-3)$$

where $Q_{(T)}$ is the flood peak in m^3/s of recurrence interval (RI) T -years, $c_{(T)}$ is the runoff coefficient for a T -year event, $i_{(T_c, T)}$ is the T -year storm rainfall intensity in mm/hr of duration equal to the time of concentration T_c (hours) of the catchment and A is the catchment area in km^2 .

In this approach, the value of $c_{(T)}$ purports to transform a T -year design storm $i_{(T_c, T)}$, of duration T_c , into a T -year flood peak $Q_{(T)}$ for a catchment of area A . The variable $i_{(T_c, T)}$ can be determined, for a particular site, from suitable Intensity-Duration-Frequency (IDF) relationships of design storms. However, the estimation of the runoff coefficient $c_{(T)}$ remains the main source of uncertainty in the probabilistic application of the rational formula. It is the least precise variable of the rational formula, in spite of it being bounded in the interval (0; 1), and suggests that a fixed ratio of peak runoff rate to rainfall rate exists for the site, which in reality is not the case (Chow et al., 1988: 497). It is the estimation of the design runoff coefficient of the rational formula that forms the main focus of a review of this method (see Chapter 4). To this end, this research investigated the calibration of the runoff coefficient, on past flood peak and flood volume pairs for a number of catchments in South Africa, to assist with its determination. The calibration of runoff coefficients on past floods is also the practice that was adopted in Australia (Institute of Engineers Australia, 1987) where it was shown that the use of calibrated coefficients in a probabilistic approach to the rational formula could consistently provide flood estimates for catchments up to $250km^2$. In this dissertation, the "data set" of runhydrographs (see Section 2.3.) produced by Hiemstra and Francis (1979) was used to calibrate the coefficients in order to investigate the probabilistic approach of the rational formula for selected catchments in South Africa.

In South African practice, the idea of calibrating the rational formula's runoff coefficient is not new. Alexander (2002) proposed a new standardised regional flood estimation technique called the standard design flood (SDF). This method is essentially a probabilistic approach to the rational formula, as advocated by Alexander (1990), utilizing calibrated runoff coefficients. The SDF method is based on the calibration of the runoff coefficient against design floods determined from a frequency analysis, using the LOG-Pearson-III (LPIII) distribution, of recorded events from a number of catchments in South Africa. According to Alexander (2002), the SDF can be applied to all sizes of catchments in South Africa, ranging in size from 10km² to 40 000km². Alexander has also suggested a standard design hydrograph for the SDF with a fixed triangular shape that has a rising limb equal to the time of concentration of the catchment T_c and a falling limb equal to $2T_c$, i.e. an effective time base-length of $3T_c$. This idealized hydrograph is the same as that proposed by Rooseboom et al., (1981) where it was noted that the runoff volume is greater than the proportionate part of the storm rainfall that runs off during the time of concentration.

In an independent test, the average ratio of Alexander's 50-year SDF flood peak to the 50-year LPIII flood peak was found to be approximately 210% (Görgens, 2002). Alexander's method was designed to be purposefully conservative and he states that the over-estimates fall within the range of uncertainties associated within all design flood procedures. However, Görgens (2002) states that although the cost and implications associated with a conscious over-design in terms of a bridge/culvert is relatively minor, by contrast it is not acceptable for dam spillway design, where the cost of the spillway is a significant component of the total dam cost. An average over-estimate of 200% might render some projects infeasible. As such, Görgens recommends that the SDF should be seen as a conservative approach similar to that of the regional maximum flood (RMF) method.

Conscious of this, the approach adopted in this investigation was slightly different in that the calibration of the runoff coefficients was performed on past flood peak and volume pairs (as offered by the runhydrograph method). It was anticipated that this would yield coefficients that could, in a design situation, describe a complete design flood hydrograph (peak, volume and time base-length). The methodology and results of this investigation are described in Chapter 4 while the theory behind the runhydrograph method is explained in Section 2.3. This particular investigation formed the basis for a journal article (Parak and Pegram, 2006) which is included in Appendix B.

2.3. The runhydrograph method

The runhydrograph method was developed by Hiemstra and Francis (1979) and was based on earlier work by Hiemstra (1972, 1973 and 1974), Hiemstra et al. (1976) and Francis (1979). It is based on the joint probability analysis of same-event flood peak and flood volume pairs of recorded data from 43 catchments in South Africa (see Table A1 in Parak and Pegram (2006) included in Appendix B). Hiemstra and Francis discovered that the natural logarithms of the flood peak and its corresponding volume were approximately normally distributed and well correlated, with a cross-correlation coefficient with mean 0.78 and standard deviation 0.12 (a relatively narrow range whose mode is 0.85).

Fig. 2-2 (from Hiemstra and Francis, 1979: 14) shows the natural logarithms of the recorded flood peak and volume pairs plotted together with the contours of equal probability density of a standardized bivariate normal probability density function (with a cross-correlation coefficient of 0.85). The correlation of the peak-volume pairs can be seen from Fig. 2-2 in that the plotted peak-volume pairs cluster around the 45° line in an elliptical shape. Also shown in Fig. 2-2 (in the positive quadrant) are 10- and 100-year return period exceedence probability contours (bold lines). The contours describe the joint probability of flood peak and flood volume exceedence. The dashed lines intersecting on the 100-year exceedence contour include an area in which the bivariate probability density function integrates to 0.01. Thus, on average, 1% of the observations will lie within this area, and within other areas of bivariate exceedence similarly defined on the 100-year contour.

It is further evident from Fig. 2-2 that the exceedence contours are able to produce “families” of hydrographs (peak-volume pairs) of equal probability of jointly being exceeded, but of varying shape. These families can range from the marginal peak (associated with any volume), to the “most likely” joint peak and volume pair through to the marginal volume, each with an equal probability of joint exceedence. However, it can also be seen from Fig. 2-2 that the plotted peak-volume pairs are very well correlated. If the cross-correlation coefficient approaches unity, the minor axis of the ellipse reduces to zero. Thus, although more than one combination of a peak-volume pair exists that has the same probability of jointly being exceeded, the most likely (modal) pair will be found at the intersection of the 45° line on the exceedence contour (the point where the probability density is highest).

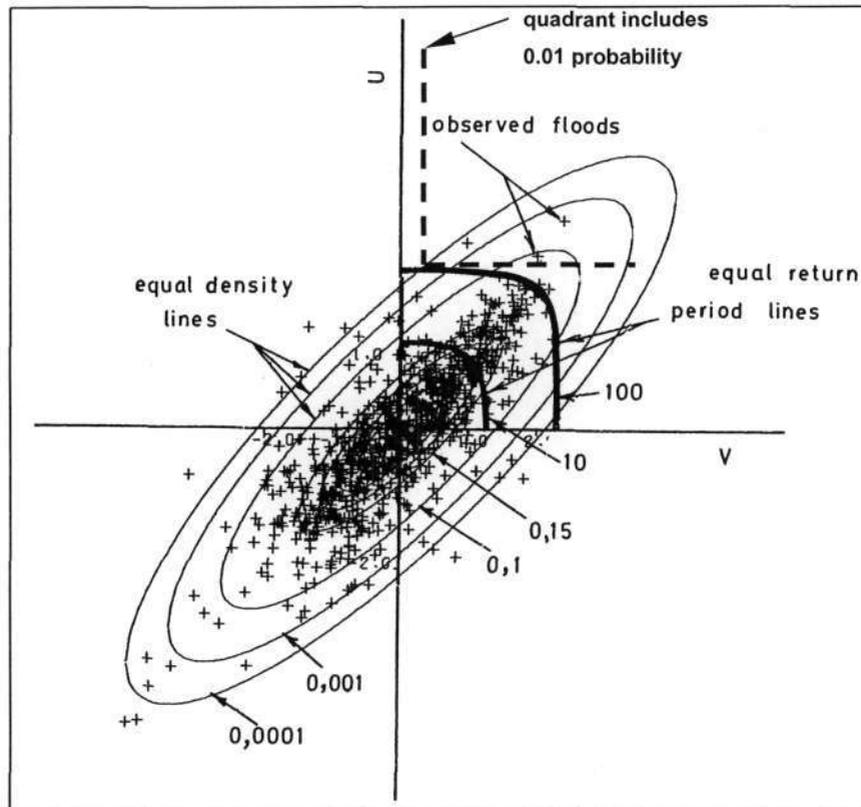


Figure 2-2. A standard bivariate normal probability density function, with a cross correlation coefficient of 0.85, plotted with log-transformed observed flood peak-volume pairs in probability space (from Hiemstra and Francis, 1979: 14). The bold lines in the positive quadrant are the 10- and 100-year return period joint-exceedence contours. The dashed lines include a quadrant to the upper right, which on average will include 1% of the observations.

Fig. 2-3 shows the application of the runhydrograph method as suggested by Hiemstra and Francis (1979) for design flood peak and volume estimation (for a cross-correlation coefficient of 0.85). The numbers listed on the top right of Fig. 2-3 are the standardized ordinates of the peak-volume exceedence contours for the selected recurrence intervals. They describe the joint exceedence of the most likely peak-volume pair (corresponding to line #1) through to the exceedence of the marginal peak (corresponding to the vertical axis to the left of line #6). However, it is unlikely that a peak-volume pair will occur on *lines 4, 5 and 6* for this relatively high correlation, and as stated above, the most likely (modal) pair will be found at the intersection of the 45° line and the exceedence contour.

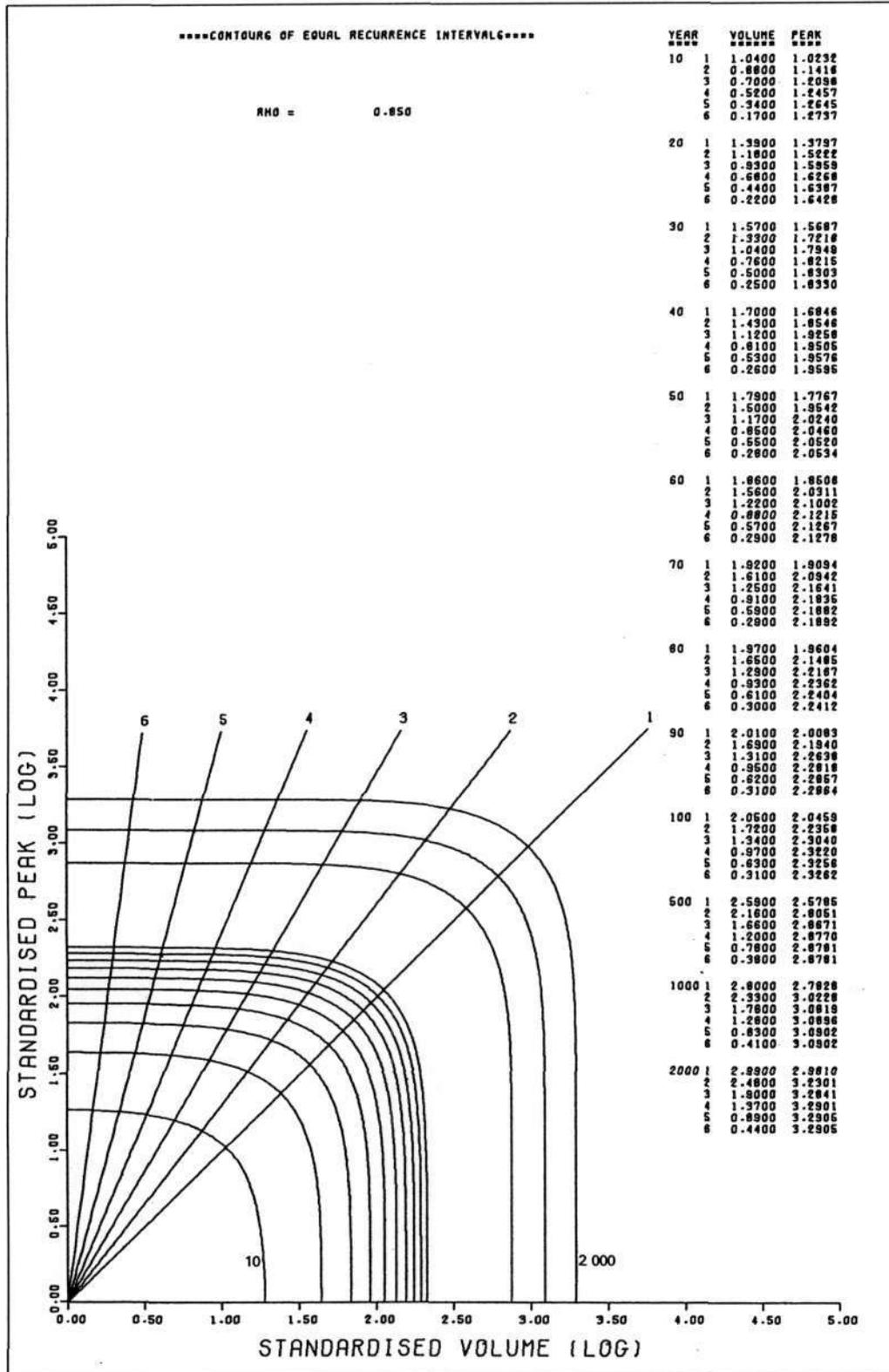


Figure 2-3. Joint flood peak and flood volume exceedence contours, in probability space for a peak-volume cross-correlation coefficient of 0.85 (from Hiemstra and Francis, 1979: 53).

The reason for choosing the runhydrograph as the data set against which to calibrate the runoff coefficients towards a probabilistic approach of the rational formula (see Section 2.2.), is that the runhydrograph method summarises for a given catchment the family of characteristic peak and volume discharges for a given recurrence interval. This method was based on the frequency analyses by Hiemstra and Francis (1979) of all rare hydrographs (which were carefully screened for reliability) in a continuous stream flow record. Furthermore, these hydrographs were analysed independently of rainfall input and catchment characteristics. Despite its apparent merit, this method is unfortunately seldom used in design situations. However, it was felt that this set of statistics would be valid to use for calibration of the rational formula's runoff coefficient for this research and the modal peak-volume pair was chosen for this purpose in order to limit the number of variables. The details of this investigation are explained in Chapter 4 and are sourced from Parak and Pegram (2006), which is included in Appendix B.

2.4. Chapter summary

Chapter 2 introduced the three approaches of design flood prediction, i.e. empirical, deterministic and probabilistic, and three methods used to this end. Common difficulties with two of the methods, namely the RMF and the rational formula methods, were also discussed. These methods are reviewed in the Chapters 3 and 4 with the intention of resolving these matters, namely:

- The association of a return period with the RMF.
- The addition of landscape parameters in improving empirically derived flood peak estimates.
- The calibration of the rational formula's runoff coefficient on runhydrograph floods towards a probabilistic implementation of this method.

CHAPTER 3

3. A REVIEW OF THE REGIONAL MAXIMUM FLOOD (RMF)

Chapter 2 introduced the RMF method of flood prediction. The RMF (Kovačs, 1988) is an empirically derived upper limit flood peak that can reasonably be expected at a given site. This flood is computed from an empirical equation, the Francou-Rodier equation (Francou and Rodier, 1967), based on the size of the catchment (area) and on the catchment's location (a regional value that is indicative of the site's flood magnitude potential). It was further discussed in Chapter 2 that, in a design situation, one is not able to easily associate an exceedence probability with RMF-based flood estimates. Kovačs (1988) estimated the return period of the RMF to be in the order of, but greater than, 200 years (Kovačs, 1988: 19). However Görgens (2002), in his review of this matter, stated that the return period of the RMF is much larger than 200 years. Thus, this method tends to be used by practitioners as an upper limit flood estimate.

This chapter describes the investigation into the estimation of a return period which could be associated with the RMF-based flood estimate by simultaneously plotting the floods determined from this method with probabilistically modelled floods for the same catchments. Furthermore, this chapter also examines the extension of empirically based methods (as a function of catchment area only) through the inclusion of other measures of the fluvial landscape. This was assessed by calibrating an empirical equation for a catchment on landscape parameters. The equation was then validated by comparing it with probabilistically modelled floods for the same catchments. Central to both investigations is the use of flood records, which were statistically modelled using the General Extreme Value (GEV) Distribution.

The study described here formed the core of a journal article (Pegram and Parak, 2004). This article also details additional investigations that were carried out in this study concerning issues related to flood prediction, of which Chapter 3 is a summary of the main outcomes. The article is included in its entirety in Appendix A.

3.1. Methodology and results

3.1.1. Flood record modelling

The flood record database consisted of annual flood peaks from 130 catchments across South Africa. This data was a subset of the actual records used by Kovačs in the construction of the RMF curves (Kovačs, 1988) and had lengths of record ranging from 9 to 76 years. This data set, although old (final year of record was 1988), provided the starting point for this study. The return period associated with each annual peak was computed using the Weibull plotting position (Weibull, 1939). This method was used as it is considered as being more conservative than the Cunnane plotting position (Cunnane, 1978); the Weibull plotting position associates a shorter return period with its highest ranked flood than the Cunnane plotting position. The Weibull plotting position is expressed as:

$$T = \frac{N+1}{r} \quad (3-1)$$

where T is the return period (in years), N is the length of record (in years) and r is the rank of the flood peak; $r = 1$ for the largest peak.

Using Eq. 3-1, the return period associated with each annual peak of a catchment was computed. Following the work of De Michele and Salvadori (2002) and Kjeldsen et al. (2002), the probability distribution of these peaks was assumed to follow a Generalized Extreme Value (GEV) Distribution (Jenkinson, 1955). This distribution has the following form:

$$Q_T = \mu + \sigma y_T \quad (3-2)$$

where Q_T is the T -year return period flood peak estimate, μ and σ are shift and scaling parameters respectively and y_T is the GEV reduced variate corresponding to a T -year return period, i.e.

$$y_T = \frac{1}{k} \left[1 - \left\{ -\ln \left(1 - \frac{1}{T} \right) \right\}^k \right] \quad (3-3)$$

where k is a shape parameter and T is the return period. When $k = 1$, the GEV Distribution reduces to the Extreme Value Type 1 (EV1) Distribution (Gumbel, 1941).

The GEV Distribution was fitted to the observed peaks, by using the return period T computed using Eq. 3-1 to calculate the reduced variate y_T , and by fitting the rest of the variables (shift, scaling and shape) by minimizing the sum of the squares of the differences between the observed and the modelled peaks. Although the probability distribution could have been fitted directly to the peaks, using the *method of moments* or the *method of maximum likelihood* (instead of using a plotting position to estimate the peaks' probability value), the method adopted in this investigation was thought to be simpler yet still valid.

This model of the flood data formed the basis with which to carry out the investigations described. Some of these data and their distribution fits are presented in Table A1 (Part 3) in Pegram and Parak (2004) included in Appendix A.

3.1.2. Return period of the RMF

The RMF was estimated for 57 catchments for which both annual flood peak data were available and modelled as explained in Section 3.1.1, and where the regional K -values were obtained from Kovačs (1988). These were 15 catchments for Region 4.6, 30 catchments for Region 5 and 12 catchments for Region 5.2. The delineation of the regions is shown in Fig. 2-1 (Kovačs, 1988). The method employed in this investigation was to coaxially plot the 50-, 100- and 200-year probabilistically modelled floods with that from the RMF (corresponding to the same regions and catchments). These were plotted against catchment area as the independent variable as shown in Fig. 3-1. Since the return periods of the modelled floods were known, the return period of the RMF could then be visually estimated. The results for Regions 5.2, 5 and 4.6 are shown in Figs. 3-1, 3-2 and 3-3 respectively, where the 200-, 100- and 50-year probabilistically modelled flood magnitudes (Q_{200} , Q_{100} and Q_{50}) are represented by the thin solid line, the dashed line and the dotted line respectively. The RMF-based floods (Q_{RMF}) are represented by the thick solid lines.

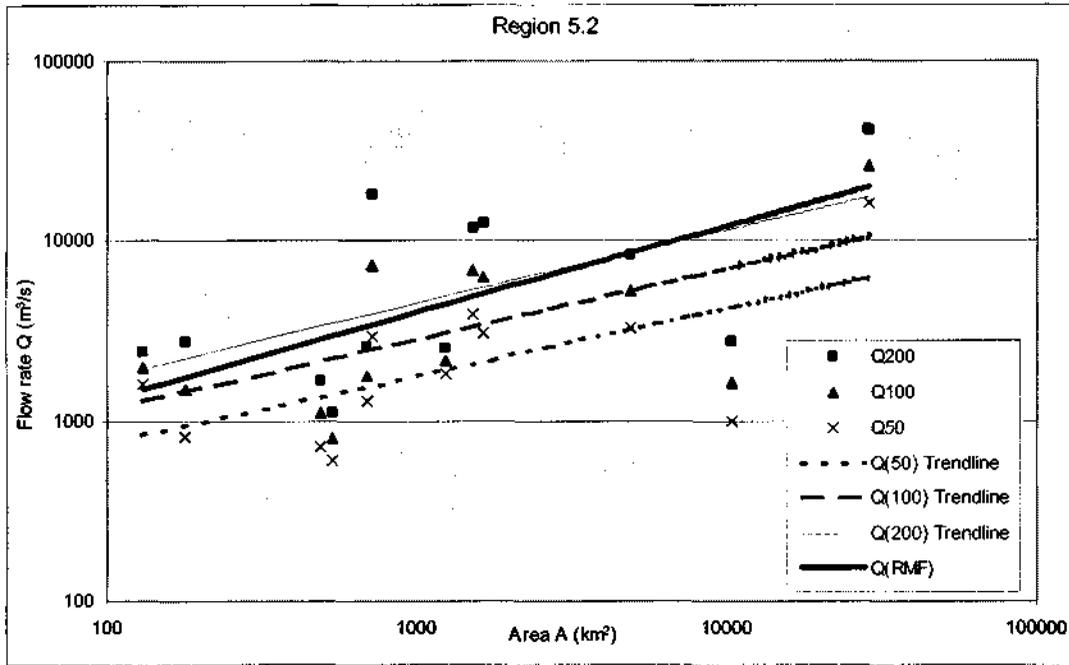


Figure 3-1. Determination of the return period associated with the RMF for Region 5.2. The bold line is the RMF estimate (defined by the Francou-Rodier equation) and the thin lines (dotted, dashed and solid) are trend-lines fitted to the 50-, 100- and 200-year floods estimated from a probabilistic analysis of recorded annual peaks for catchments in Region 5.2.

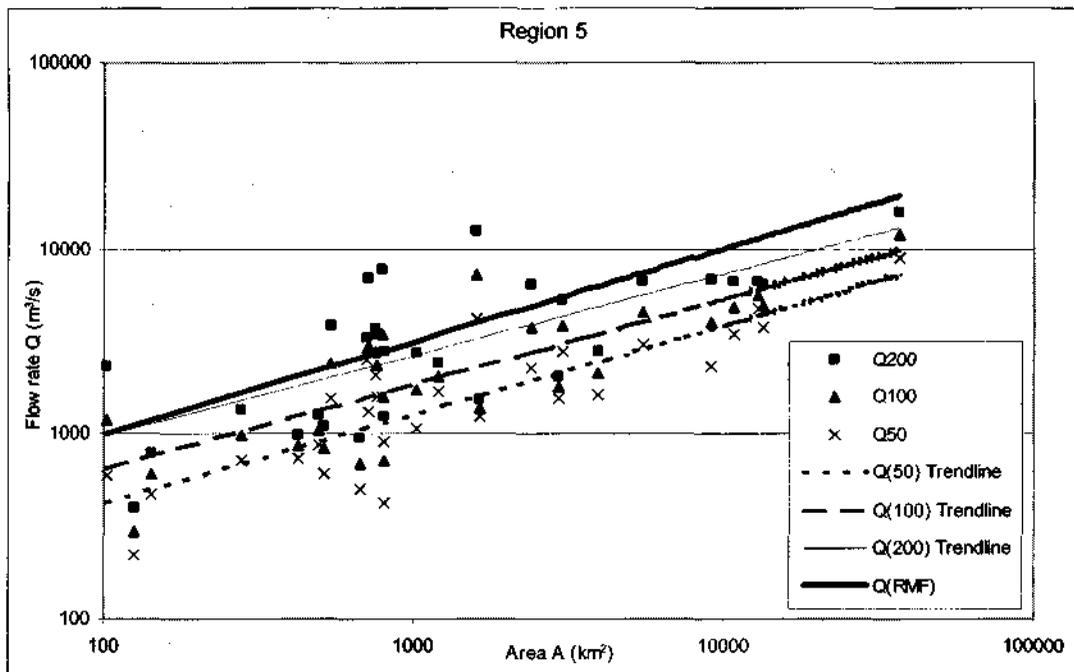


Figure 3.2. Determination of the return period associated with the RMF for Region 5 (description as per Fig. 3-1).

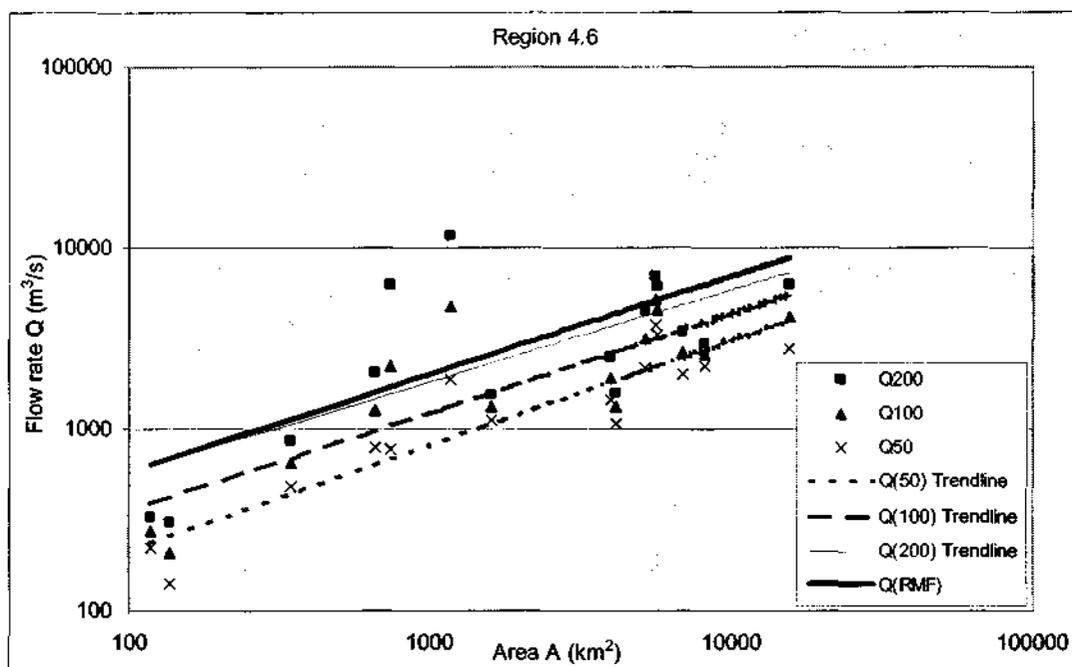


Figure 3-3. Determination of the return period associated with the RMF for Region 4.6 (description as per Fig. 3-1).

From Figures 3-1 to 3-3, it is clear that the RMF is closest to the 200-year flood when compared to the 50-, 100- and 200-year probabilistic flows. Table 3-1 summarizes the trend-line equations and regressions from Figs. 3-1, 3-2 and 3-3. It can be seen in Table 3-1 that the 200-, 100- and 50-year trend-line slopes are slightly flatter than the RMF curves for all the floods in all the regions except one (Q_{50} for Region 4.6). The coefficients of determination (R^2 -values) range from poor (0.34) to good (0.84), appearing to improve for the more frequent floods and for the regions located further inland (Region 4.6). However, despite some poor fits, the correspondence is generally fair to good and provides an indication of the approximate magnitude of the RMF, i.e. it is closely approximated in all three regions by the 200-year flood.

Table 3-1. Summary of the trend-lines from Figs. 3-1, 3-2 and 3-3. The coefficients of determination (R^2 -values) appears in parentheses.

	Q_{RMF}	Q_{200}	Q_{100}	Q_{50}
Region 5.2	$145A^{0.48}$	$269A^{0.41}$ (0.34)	$191A^{0.39}$ (0.38)	$134A^{0.38}$ (0.41)
Region 5	$100A^{0.5}$	$129A^{0.44}$ (0.51)	$77A^{0.46}$ (0.63)	$45A^{0.48}$ (0.71)
Region 4.6	$48A^{0.54}$	$55A^{0.51}$ (0.51)	$29A^{0.54}$ (0.70)	$20A^{0.58}$ (0.84)

As an extension to this investigation, an attempt was made to examine if the k -parameter of the GEV Distribution could be regionalised using the regional description of Kovač's (1988) K -values. Despite the potential merit of this attempt, the results were poor and there did not appear to be any relationship between the two parameters. This result is shown in Fig. 3-4 where the k -parameter of the GEV Distribution for each catchment is sorted using the RMF K -value for the catchments. Fig. 3-4 also shows the standard errors associated with the mean k -parameter for each regional K . Although for Region 5 the least error is observed, it must be noted that this region has the most data. However the spread of the k -parameters in each region is large and the results were not successful.

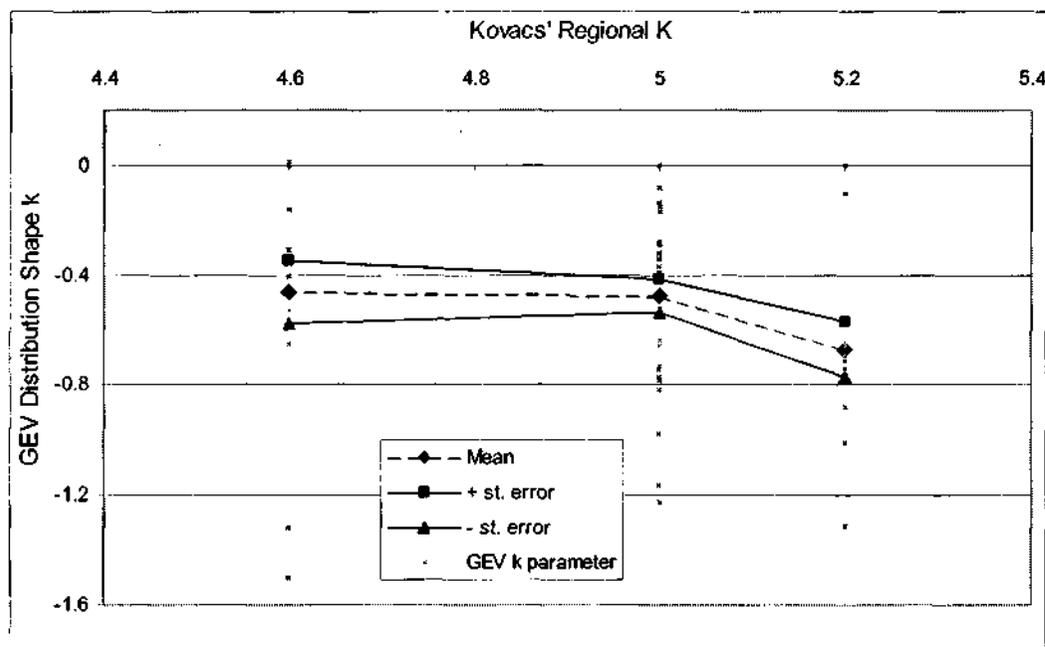


Figure 3-4. Plot of the k -parameter of the GEV Distribution for each catchment against the regional K -value of the RMF (Kovač, 1988) for Regions 4.6, 5 and 5.2. The mean GEV k -parameter together with its standard errors is also shown for each regional K .

3.1.3. Inclusion of landscape parameters in empirical formulas

In this investigation an attempt was made to determine if other measures of the landscape besides catchment area (such as mean channel slope, mean annual precipitation, drainage density, catchment relief and ruggedness number) were able to improve empirical equations based on *catchment area* only. In order to examine this, an empirical equation based on *area* and some other landscape measure(s) was calibrated on the probabilistic flows, modelled as explained in Section 3.1.1, for certain catchments (calibration catchments). In order to validate the calibrated empirical

equation, it was used to generate design flows for catchments not used in the calibration exercise (validation catchments). These flows were then compared to probabilistically modelled flows (see Section 3.1.1) of the validation catchments. The extent to which the generated flows (from the empirical equations) mimicked the probabilistic flows was adjudged on the strength of the coefficient of determination (R^2 -value). The success, or lack thereof, of a calibrated empirical equation including landscape parameters in addition to *catchment area*, was assessed on the strength of R^2 in validation when compared to this value (R^2) when an empirical equation calibrated on *catchment area* only was used. Thus, through a series of step-wise regressions it could be determined if an empirical equation calibrated on *catchment area* only was improved through the inclusion of additional landscape data.

To this end, landscape data were used from a study by Parak (2003) for 25 catchments for which the peak discharges of the catchments were probabilistically modelled as explained in Section 3.1.1. In his study, Parak (2003) captured morphometric data for catchments across the country in an investigation into the relationship between floods and landscape. Parak (2003) used already catalogued data, such as Petras and Du Plessis (1987) and Kovačs (1988) and supplemented this with further data through map work accessed electronically from Midgley et al. (1994). As mentioned earlier, the following landscape data were used in this investigation: catchment area, mean channel slope, mean annual precipitation (MAP), drainage density, catchment relief and ruggedness number. These are listed in Table A1 in Pegram and Parak (2004), which is attached in Appendix A. Typical catchment morphometry and its derived geometry are shown in Fig. 3-5 (from Parak, 2003). Other parameters (such as the various length measures of water courses and catchment perimeter) were not incorporated in the formulation of the empirical equations since they are highly correlated with *catchment area* (being indicative of size). Thus it was felt that these parameters would not add more information compared to using *catchment area* on its own, hence they were omitted.

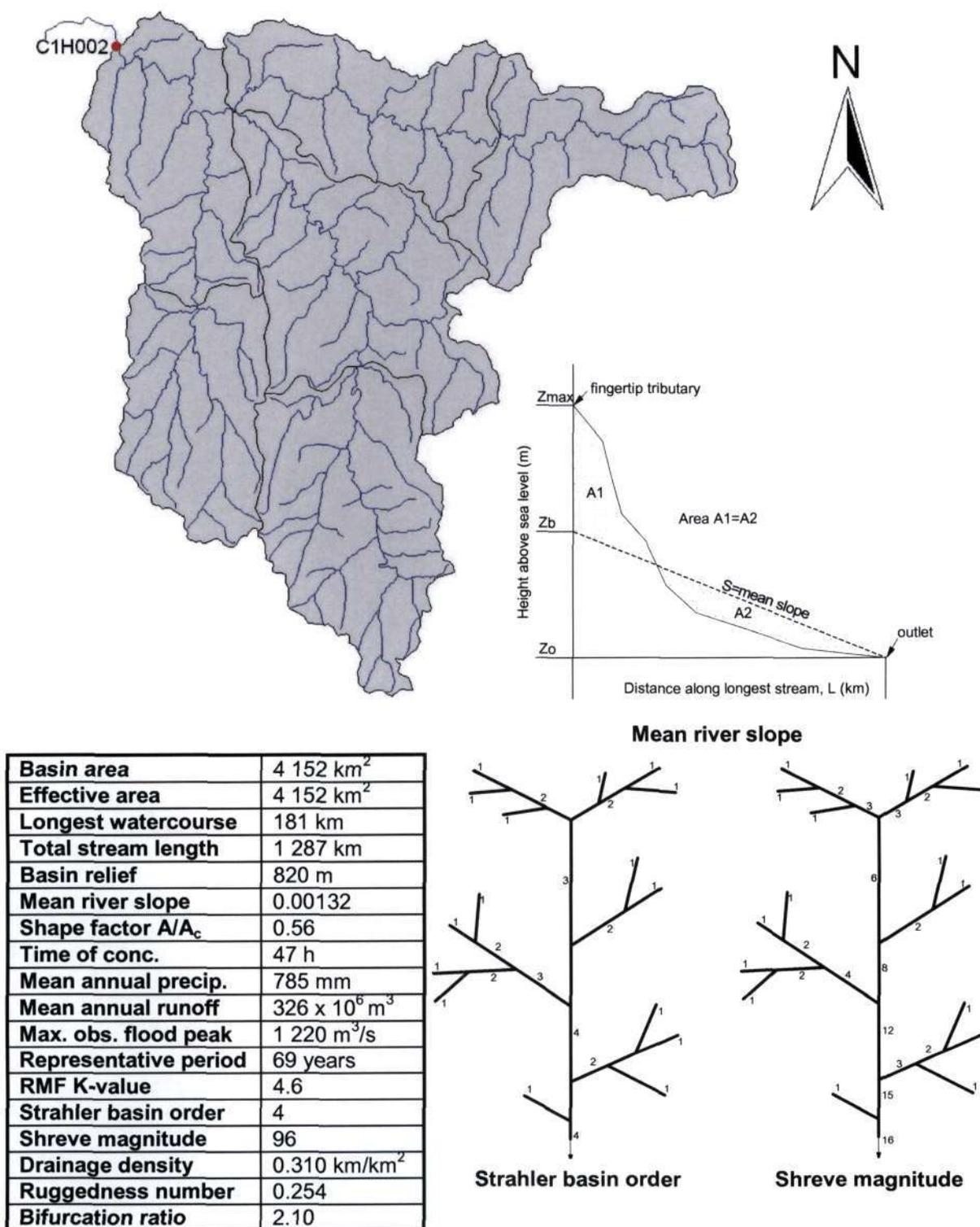


Figure 3-5. The Klip River catchment (represented by gauge C1H002) and basin properties (derived from Petras & Du Plessis, 1987; Kovačs, 1988; Midgley *et al.*, 1990 and Parak, 2003). The definitions of the *mean river slope*, *Strahler basin order* and *Shreve magnitude* are also shown. Reference should be made to Pegram and Parak (2004) in Appendix A for the definitions of these and other parameters.

The flood and landscape data were randomly split into two groups, one for calibration and the other for validation. Group 1 and 2 consisted of 13 and 12 catchments respectively. The roles of these groups were also swapped, meaning that in the first test Group 1 was used for calibration and Group 2 for validation, while in the second test Group 2 was used for calibration and Group 1 for validation. The magnitude of the flow rate that was used for calibration and validation was the 20-year flood because all the flow records used here had observation periods greater than 20-years. It was also argued that this flood would be the least likely estimate to be affected by fitting the wrong probability distribution.

It is important to note that landscape data are sensitive to map scale; i.e. different values of the parameters will be obtained at different scales. For example, the river detail shown on a larger scaled map is much less than that which is shown on fine-scaled maps. This has a direct influence on the magnitudes of the landscape parameters. Measures such as total stream length, stream orders, drainage densities and ruggedness numbers are all dependent on the scale of the map from which these parameters were extracted. However, the use of finer scaled maps comes at the expense of greater effort and time requirements for data extraction. In order to minimise comparative errors, Parak (2003) used uniform scaled maps (at 1:250 000) from Midgley *et al.* (1990) to extract data that are sensitive to scale.

In Parak's study (2003), the most suitable formulation of an empirically based flood-landscape equation was discovered to be a power-law relationship after studying various literature, in particular Patton (1988). This took the form of:

$$Q_{20} = aA^b X^c Y^d \dots \quad (3-4)$$

where Q_{20} is the 20-year flood (used in this instance), A , X and Y are landscape quantities and a , b , c and d are parameters to be regressed from the data. The formulation for the regression equation was to take logarithms of Eq. (3-4) and regress using the linearised model:

$$\log(Q_{20}) = \log(a) + b.\log(A) + c.\log(X) + d.\log(Y) \quad (3-5)$$

Fig. 3-6 shows the calibration of an empirical equation defining the 20-year flood (Q_{20}) as a function of catchment area only. The R^2 statistic implies a strong relationship

(0.86) and good fit. When this model ($Q_{20} = 24.6A^{0.517}$) is tested against the data reserved for validation (shown Fig. 3-7), the fit is poor (producing a moderate R^2 statistic of 0.54). These two statistics of R^2 , 0.86 in calibration and 0.54 in validation, then became the target set against which to compare the improvement (or lack thereof) of the empirical model through the inclusion of further landscape parameters. These statistics (of the *step-wise* regressions) are summarised in Table 3-2.

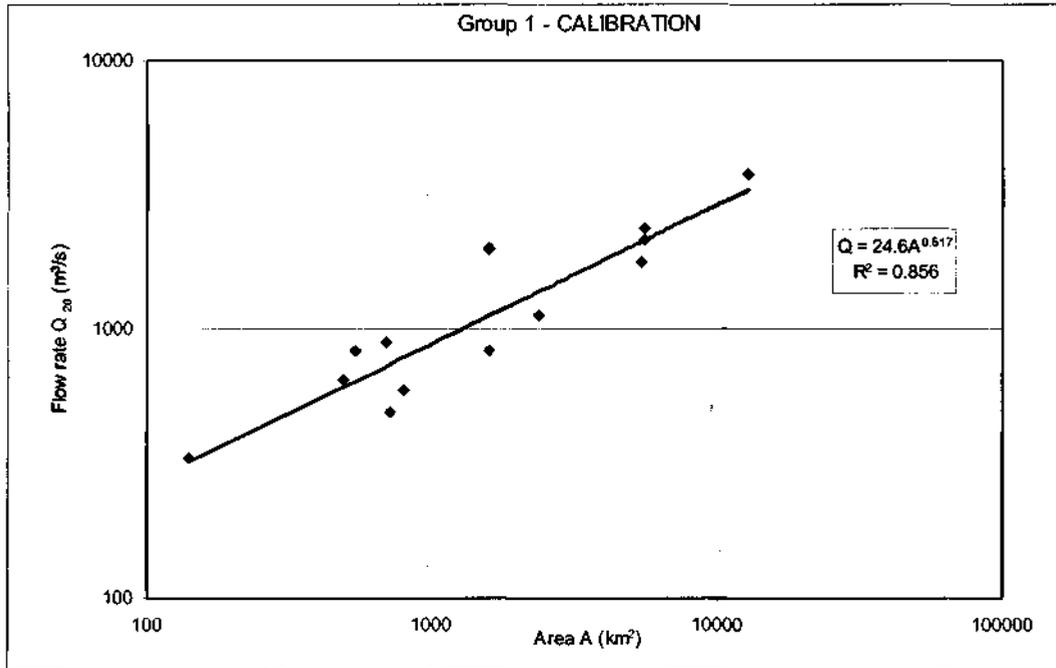


Figure 3-6. The 20-year flood (Q_{20}) vs. *catchment* area for the calibration of an empirical model. The number of catchments used here was 13.

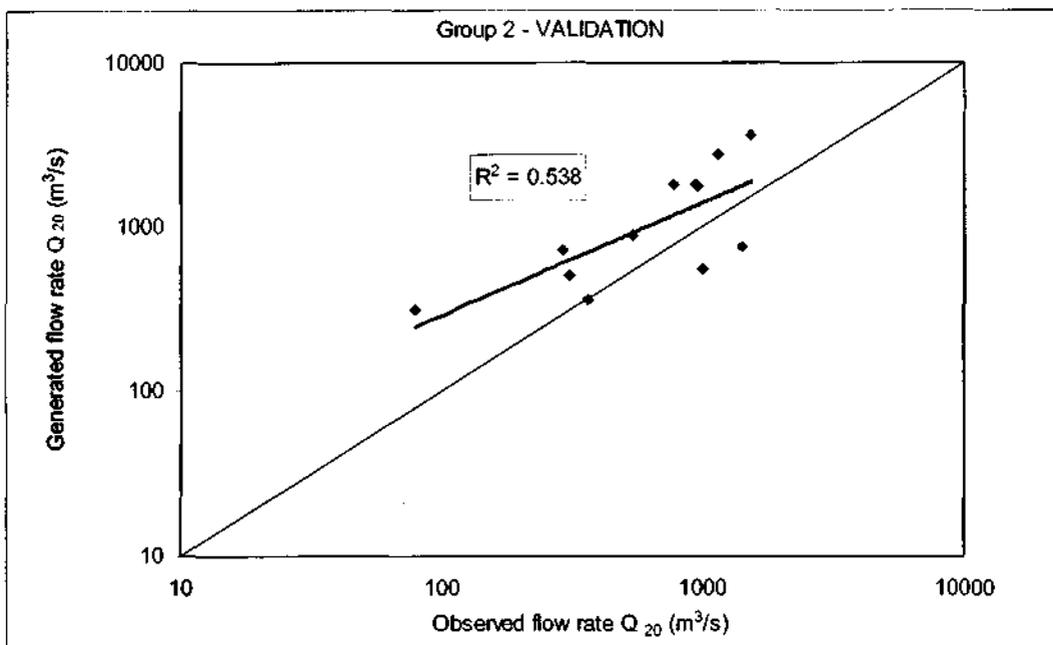


Figure 3-7. The 20-year floods generated from the model calibrated in Fig. 3-6 vs. the probabilistically modelled floods for the same catchments. The number of catchments used in this validation was 12.

Table 3-2: Results of the step-wise regression of the model calibration and validation. The numbers in parentheses flag the ranked "best fit" (based on the R²-value) to the validation data.

		R ²	
		Group 1: calib. Group 2: valid.	Group 2: calib. Group 1: valid.
Q ₂₀ vs..	Area	Calibration: 0.86 Validation: 0.54 (3)	0.54 0.86 (1)
	Area and slope	Calibration: 0.87 Validation: 0.53	0.57 0.72
	Area and MAP	Calibration: 0.89 Validation: 0.51	0.55 0.77 (3)
	Area and drainage density	Calibration: 0.87 Validation: 0.53	0.54 0.78 (2)
	Area and relief	Calibration: 0.88 Validation: 0.56(1)	0.64 0.63
	Area and ruggedness number	Calibration: 0.88 Validation: 0.55	0.59 0.66
	Area, slope and MAP	Calibration: 0.90 Validation: 0.50	0.64 0.39
	Area, ruggedness number and MAP	Calibration: 0.92 Validation: 0.52	0.60 0.64
	Area, drainage density and MAP	Calibration: 0.89 Validation: 0.49	0.57 0.53
	Area, relief and MAP	Calibration: 0.89 Validation: 0.55 (2)	0.65 0.61

It is evident from Table 3-2 that the addition of landscape data as independent variables in the prediction of floods does not appear to improve the flood prediction

ability of the empirical model. This is drawn from the fact that the R^2 -value does not appear to significantly improve in validation of the empirical models (calibrated on *area* and some other landscape measure) when compared to the *area*-based empirical model (calibrated on *area* only). For example, when Group 1 was used as the calibration set and Group 2 the validation set, the R^2 -values in calibration and validation of the *area*-based model were 0.86 and 0.54 respectively. Thereafter, when including other measures of the landscape in addition to *area* in the calibration and validation of the models, the predictive ability of the additional landscape parameters was witnessed to be small. This is evident in that the R^2 -values in calibration and validation of the model including additional landscape parameters did not improve significantly when compared to 0.84 and 0.54 respectively (for the *area*-based model). The best overall improvement (although slight) in R^2 for calibration and validation of an empirical model was when *relief* was included with *area* (0.88 and 0.56 respectively). In spite of a relatively high value of R^2 experienced in calibration of a model including *ruggedness number* and *MAP* with *area* (0.92), the R^2 -value in validation was poorer (0.49) than when *area* was used alone. When the roles of the groups were reversed, i.e. Group 2 was used for calibration of the models and Group 1 for validation, the R^2 -values for calibration of the empirical models were poorer than this statistic in validation. Despite this, it was still evident that models calibrated with *area* and some other landscape measure fared poorer in validation than with models calibrated with *area* only. Thus for both tests, i.e. for the first test where Group 1 was the calibration set and Group 2 the validation set, and for the second test where Group 2 was the calibration set and Group 1 the validation set, it seems that the best model of floods and landscape is simply *area*-based. The probable reason for the fact that Group 1 seems stronger in calibration and validation than Group 2 is possibly due to the small sizes of the groups (respectively 13 and 12 catchments) bearing in mind that the groupings were a random choice process.

3.2. Discussion of results

3.2.1. Return period of the RMF

Probabilistically modelled and RMF flood estimates were plotted in Figs. 3-1, 3-2 and 3-3 against catchment area for three *K*-regions in order to estimate a return period that could be associated with the RMF. It is evident in all three figures that the RMF curve and the trend-line fitted to the 200-year return period flows closely approximate each

other. However, mindful of the contents of Table 3-1, the lines are not parallel and, for some of the regions, the trendlines have a poor fit with scatter observed, especially for catchments with areas in the order of 1000km². On further investigation, it turns out that the outlying 200-, 100-, and 50-year probabilistic floods were skewed by large flood peaks that were observed in a relatively short return period (between 20 and 30 years) for those catchments. Although this problem is not addressed here, it is expected that with more data, the effect of the outliers will be diminished.

The result is that the plot for Region 5 (Fig. 3-2) is likely to be more representative of the relationships than Figs. 3-1 and 3-3 as it contains more data. Referring to Fig. 3-2 and Table 3-1, the R^2 -values for the trend-line fits of the Q_{200} , Q_{100} and Q_{50} flows range from fair (0.51) to good (0.71) respectively. Although for the largest catchment, the RMF-based flood estimates appear in Fig. 3-2 to be greater than the Q_{200} trend-line by a factor of about 1.5, the estimates are all within the same order of magnitude. Based on the results of this comparison for all three regions, it would be reasonable for design purposes to assume the RMF to have a return period of the order of 200 years. To ascribe longer return periods might cause "under-design" to become prevalent.

3.2.2. Inclusion of landscape parameters in empirical formulas

The improvement of area-based empirical equations, through the inclusion of other measures of the landscape, was examined by a series of step-wise regressions. Firstly an empirical equation based on area was calibrated. This model was then validated against probabilistically modelled flows. Further models were then calibrated and validated, this time with the inclusion of landscape parameters. The improvement of the area-based model was then examined based on the improvement of the R^2 -value when in calibration and validation of the new model. Table 3-2 summarises the R^2 -values for these tests.

However, it must be noted that the effect of regionalization was ignored in this exercise and all data used (25 stations) were pooled before random selection of the two groups. The attempt to differentiate them by using their geomorphological characteristics via step-wise regressions of the log of a power-law equation (Eq. 3.5) did not appear to have any apparent skill (see Table 3.2).

The conclusion that can be drawn from these tests is that the inclusion of landscape parameters in addition to *catchment area* does not appear to improve the predictive ability of empirical models when compared to the use of *catchment area* as the sole independent variable of such models. This result is based on the negligible improvement of the R^2 -value in calibration and validation when additional landscape data are added to *catchment area*, as evidenced in Table 3-2. The results of this exercise, based on the data available here, imply that the role of landscape in flood production is minor and that the inclusion of landscape measures does not materially improve flood prediction. It would seem that the use of *catchment area* is the simplest and best of all measures in estimating flood magnitudes from empirical methods.

3.3. Chapter Summary

Chapter 3 investigated two issues with regard to empirically-based methods, namely the exceedence probability of the RMF estimate and whether the inclusion of other measures of the landscape (in addition to area) assist in improving the flood prediction capabilities of *area*-based empirical models. In the first part of the investigation, it was discovered that the 200-year probabilistically modelled flood closely approximates the RMF-based flood for three *K*-regions tested. Thus it is concluded that in a design situation, it would not be unreasonable to take the return period of the RMF as 200 years. In the second part of the investigation, it was discovered that the use of an *area*-based model in estimating empirically-derived flood magnitudes provides sufficiently good results and that the inclusion of other measures of the landscape do not improve this model.

CHAPTER 4

4. THE RATIONAL FORMULA FROM THE RUNHYDROGRAPH

In Chapter 2 attention was drawn to the fact that the runoff coefficient (c) of the rational formula is the least precisely defined parameter of this method. In almost all design situations, the estimation of the coefficient is subjective and left up to the experience and judgement of the designer. As a consequence, this part of the dissertation focuses on the calibration of the runoff coefficient on past flood records, following the approaches used by the Institute of Engineers Australia (1987) and Alexander (2002). However in this instance, the initiative was taken to calibrate the coefficients on a catchment's characteristic design flood peak and volume discharge (which is independent of rainfall) for catchments in South Africa. This latter resource is offered by the runhydrograph method of Hiemstra and Francis (1979), which was also introduced in Chapter 2. It was expected that, in following this route, conclusions in respect of c could be drawn from comparing the rational formula with a method that is independent of rainfall.

Chapter 4 presents the methodology and results of this research. This investigation resulted in the publication of a journal article (Parak and Pegram, 2006) and is included in Appendix B.

4.1. Methodology and results

The methods employed in this investigation were typical of those used in the derivation of a probabilistic rational formula utilizing calibrated coefficients. The explanation that follows is adapted from Pilgrim and Cordery (1992):

- *Where a set of long and reliable record of flood data from a particular catchment exists, a frequency analysis should be carried out on the observed data to determine design values of flood peaks for a range of recurrence intervals. In this study, T -year flood peak and volume pairs (Q_T in m^3/s and V_T in m^3 respectively) for the modal runhydrograph flood was computed for each of the selected catchments for return periods of 10-, 20-, 50-, 100- and 200-years. These appear in Tables A2 to A6 in Parak and Pegram (2006) in Appendix B.*

As a result of this, values of B , the time base-length of the triangular approximated hydrographs, were also computed.

- A design formula for the calculation of time of concentration T_c must be selected and used consistently throughout the derivation and use of this method. In this study the Kirpich (1940) formula was used, following the lead of Petras and du Plessis (1987):

$$T_c = 0.0633[L^2 / S]^{0.385} \quad (4-1)$$

where T_c is the catchment's time of concentration (in hours), L is the length (in km) of the longest water course and S is the slope of the longest water course.

- Design rainfall intensities, $i_{(T_c, T)}$, for the corresponding time of concentration of the catchment and recurrence interval, should be determined from a suitable Intensity-Duration-Frequency (IDF) database. These were determined from Smithers and Schulze's (2003)¹ design rainfall data-base for South Africa. These data also appear in Tables A2 to A6 in Parak and Pegram (2006) included in Appendix B.
- From the design flood peak and design rainfall data, values of $c_{(T)}$ (calibrated runoff coefficients) can be back calculated by the following equation (adapted from Eq. 2-3 in Chapter 2):

$$c_{(T)} = \frac{3.6 \cdot Q_T}{i_{T_c, T} \cdot A} \quad (4-2)$$

where $c_{(T)}$ is the calibrated runoff coefficient, Q_T is the T -year design flood peak in m^3/s (determined from the runhydrograph), $i_{T_c, T}$ is the T -year design storm rainfall intensity (in mm/h) corresponding to the catchments time of concentration T_c and A is the area of the catchment (in km^2). Values of $c_{(T)}$ determined in this way appear in Tables A2 to A6 in Parak and Pegram (2006) in Appendix B.

¹ A computer programme with a graphical user interface has been developed with this database by Smithers and Schulze (2003) to obtain design rainfall depths for any location in South Africa. The software may be downloaded from the following website: <http://www.beeh.unp.ac.za/hydrorisk/> and follow the "Design Rainfall" option.

- *The calibrated values of $c_{(T)}$ can then be regressed on any physical characteristic of the catchment.* In order to validate the calibrated coefficients at ungauged sites, regional parameters with which to relate $c_{(T)}$ with return period were sought. However, it is noted in Pilgrim and Cordery (1992) that the probabilistic runoff coefficients determined for Australia did not show much sensitivity to physical characteristics of a catchment. Mindful of this caution, South African data were used in an attempt to find a relationship.

It is important to note that the values of $c_{(T)}$ obtained in this manner are conditioned on the use of a consistent formula for the calculation of T_c and a consistent database for the derivation of the IDF rainfall relationships. A detailed explanation of each of the steps listed above is given in the following sub-sections as well as the results of each exercise.

4.1.1. Streamflow database

The streamflow database was sourced from Hiemstra and Francis (1979). The statistics of the 43 catchments that were used by Hiemstra and Francis in their study are listed in Table A1 in Parak and Pegram (2006) in Appendix B. As a point of departure, runhydrograph data from this resource were combined with catchment parameters from Petras and du Plessis (1987), namely area (A) and time of concentration T_c (based on the Kirpich (1940) formula). The number of catchments from the Hiemstra and Francis database, for which T_c values were available from the Petras and du Plessis catalogue, reduced the number of available catchments for calibration of the runoff coefficients to 29. These are listed in Table A2 in Appendix B and formed the core data set on which the rational formula calculations were performed.

4.1.2. Rainfall database

For each of the 29 catchments, a number of locations (depending on the size of the catchment) were chosen along the main watercourse for which design rainfall depths were obtained from Smithers and Schulze (2003). The output from this rainfall database provides point rainfall depths (in mm) for durations ranging from 5 minutes to 7 days and for return periods ranging from 2 to 200 years at a spatial resolution of 1 arc minute in South Africa. The mean depth for each catchment was computed and thereafter the intensity, duration and frequency (IDF) relationships were computed by

fitting a simple power-law function of storm duration to the mean rainfall depths, following Pegram (2003a). For the selected recurrence intervals, these took the form of:

- P (rainfall depth in mm) = ad^b
- i (rainfall intensity in mm/hr) = ad^c

where d is the storm duration in hours and a , b and c (which equals $b-1$) are the fitted power-law parameters. The mean intensity, corresponding to the time of concentration T_c for each catchment, was calculated from the power-law IDF relationships for the 10-, 20-, 50-, 100- and 200-year recurrence intervals – an example appears in Fig. 1 in Appendix A. The parameters fitted to the rainfall duration, for the selected recurrence intervals, are listed in Tables A2 to A8 in Appendix B. It was found that rainfall depth scaled, on average, to the power of 0.238 of rainfall duration and thus rainfall intensity to the power of -0.762 of rainfall duration with a standard deviation of 0.0419.

Area reduction factors (ARFs) were not used in this study to scale the point rainfall depths into average depths over the catchment. Instead simple averages of rainfall depths along a few points on the main watercourse within the catchment were used to account for the variation in precipitation with position and altitude for large catchments. ARFs were deemed not necessary based on the findings of Pegram (2003a), who investigated the scaling properties of rainfall in South Africa and found that they could be expressed as a function of three factors: the median one-day rainfall (which is a function of location), a function of return period (the reduced variate of the general extreme value (GEV) distribution) and a function of duration. He used this finding to modify the intensity expression of the rational formula. The storm duration used by Pegram was the catchment's time of concentration T_c determined (as in this study) from the Kirpich (1940) formula. When this duration T_c was plotted against catchment area, it was found that the points clustered about a curve to which a power-law relationship could be fitted. This was superimposed on the area reduction factor (ARF) diagram, published in the Flood Studies Report (FSR, 1975). Pegram (2003a) found that the Area vs. T_c curve yielded an almost constant ARF value of 87% across the FSR curve – see Fig. 2 in Appendix A. The implication of this is that, as long as the precipitation intensity used in the rational formula corresponds to the time of concentration of the catchment, the point rainfall is automatically scaled by a constant ARF. It is likely that the FSR's ARF curves over-estimate the relationship in South Africa, but the degree is likely to be due to climate (Pegram and Parak, 2004). Thus it

is possible that the scaling behaviour will be maintained. However, in this case the reduction factor would also automatically be absorbed into the fitted $c_{(T)}$ -values. Therefore, because c is explicitly a function of T_c , it is therefore implicitly independent of the ARF.

4.1.3. Calibration of the runoff coefficients

Based on Eq. 4-2, calibrated runoff coefficients were calculated for the 10-, 20-, 50-, 100- and 200-year return periods for each of the 29 catchments. The rainfall and flow data used were determined as explained in Sections 4.1.1 and 4.1.2. Table 4-1 contains a summary of the results. The complete set of results are given in Tables A2 to A6 in Appendix B.

The results showed that coefficients from 6 of the 29 catchments (marked with an asterisk in column 1 of Table 4-1) produced results that did not increase in magnitude with recurrence interval. As mentioned in Alexander (1990), an increase in c with return period is necessary to accommodate the known effects which also increase with return period but are not accounted for in the formula's calculation process. The main effect, requiring this increase of c with return period, is that the catchment is likely to be more saturated at the start of a storm with a longer recurrence interval (Rooseboom et al., 1981). This initial saturation, caused by pre-event rainfall, is the main reason why one can expect to obtain a higher percentage runoff with an increase in the recurrence interval of an event. Alexander (2002) states that in many of the destructive events observed, severe rainfall events were often preceded by above-normal seasonal rainfall.

Table 4-1. The results of the calibration of the c -coefficient of the rational formula on flood peaks from Hiemstra and Francis (1979). Catchments marked with an asterisk produced coefficients which decreased with an increase in return period.

No.	Station	River	Lat. (deg. dec.)	Long. (deg. dec.)	Catch. Area (km ²)	Time of Conc. T _c (h)	Calibrated c -coefficients				
							10-year	20-year	50-year	100-year	200-year
1	A2M03	Hex	25.77	27.28	494	6.4	0.301	0.303	0.304	0.305	0.306
2	A2M12	Krokodil	25.82	27.92	2 586	18	0.089	0.093	0.095	0.097	0.098
3	A3M01	Klien Marico	25.53	26.10	1 002	8.7	0.084	0.092	0.104	0.113	0.123
4	B2M01	Bronkhorst.	25.80	28.77	1 585	18.1	0.210	0.228	0.244	0.254	0.262
5	B4M03	Steelpoort	25.02	29.53	2 271	19.6	0.091	0.102	0.112	0.125	0.135
6*	B7M04	Klaserie	24.55	31.03	130	3.7	0.234	0.233	0.214	0.227	0.224
7	C1M01	Vaal	26.95	29.27	8 254	74	0.396	0.419	0.444	0.460	0.476
8	C4M01	Groot Vet	28.48	26.67	5 504	34	0.368	0.386	0.409	0.425	0.442
9*	C4M02	Vet	27.85	25.90	17 550	111	0.179	0.175	0.170	0.167	0.164
10	C5M03	Modder	29.17	26.58	1 650	18.3	0.419	0.440	0.458	0.469	0.479
11	C5M04	Modder	28.85	26.18	5 012	38	0.528	0.592	0.660	0.706	0.749
12	C5M12	Riet	29.65	25.98	2 383	23	0.218	0.235	0.252	0.264	0.274
13	C5M15	Modder	28.80	26.10	6 545	43	0.280	0.302	0.325	0.341	0.355
14	C7M01	Renoster	27.27	27.18	5 255	57	0.236	0.300	0.379	0.438	0.498
15	D1M05	Oranje	30.03	28.50	10 891	60	0.261	0.266	0.270	0.272	0.274
16*	D5M01	Renoster	31.65	20.62	2 129	27	0.263	0.264	0.264	0.264	0.264
17*	D5M04	Sak	31.65	21.77	5 799	28	0.130	0.128	0.125	0.123	0.121
18	E2M02	Doring	32.50	19.53	5 778	30	0.389	0.420	0.459	0.487	0.516
19	H1M06	Bree	33.42	19.27	754	7.6	0.454	0.457	0.461	0.464	0.468
20*	H1M07	Wit	33.57	19.15	83	2.4	0.814	0.800	0.790	0.787	0.786
21	H7M04	Huis	33.92	20.72	26	2.3	0.278	0.307	0.336	0.353	0.368
22	J2M03	Gamka	33.53	21.65	17 941	42	0.076	0.082	0.090	0.095	0.099
23	J3M04	Olifants	33.48	23.03	4 330	23	0.163	0.180	0.194	0.200	0.205
24	Q1M01	Groot Vis	31.90	25.48	9 150	18	0.089	0.097	0.108	0.116	0.124
25	Q9M10	Groot Vis	33.22	26.87	29 376	108	0.176	0.227	0.282	0.318	0.349
26	Q9M12	Groot Vis	33.10	26.45	23 041	85	0.113	0.133	0.158	0.178	0.198
27*	T3M02	Kinira	30.48	28.62	2 100	26	0.186	0.172	0.156	0.145	0.135
28	W4A03	Pongola	27.42	31.52	5 843	31	0.267	0.278	0.284	0.285	0.284
29	W5M05	Hlelo	26.83	30.73	751	17.8	0.177	0.193	0.212	0.225	0.237

Values of $c_{(T)}$, for all 29 catchments, were then coaxially plotted with c -values from Chow et al. (1988: 498) against return period in order to compare the coefficients achieved in this study. This relationship is shown in Fig. 4-1, where the coefficients from Chow et al. correspond to the "flat" slopes type (i.e. for ground slopes between 0 and 2%, since all the test catchments in this calibration exercise had slopes of less than 2%). The values from Chow et al. (1988) are also for the three "undeveloped" (rural) coverage types (i.e. cultivated land, pasture/range and forest/woodland). The coefficients from Chow et al. (1988) are shown in Table 4-2 and were determined for small rural catchments (i.e. less than 100km²) of Austin, Texas (USA).

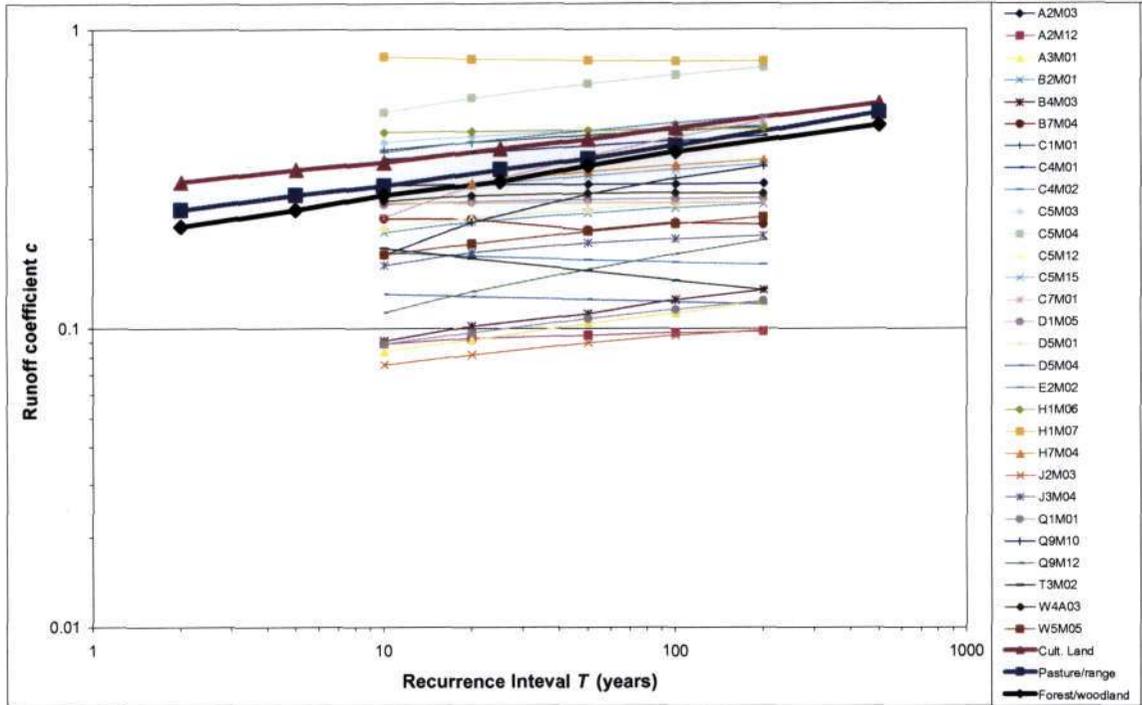


Figure 4-1. A comparison of the runoff coefficients c from Chow et al. (1988: 498) with those calibrated in this study $c_{(T)}$. The c -values plotted from Chow et al. are shown in thick bold lines and extend from the 2- to 500-year recurrence intervals.

Table 4-2. Runoff coefficients for use in the rational method for undeveloped (rural) regions in Austin, Texas in the USA (from Chow et al., 1988: 498).

Character of Surface	Runoff coefficients c							
	2-year	5-year	10-year	25-year	50-year	100-year	200-year (interpolated)	500-year
Undeveloped								
<i>Cultivated Land</i>								
Flat, 0 - 2%	0.31	0.34	0.36	0.40	0.43	0.47	0.51	0.57
Average, 2 - 7%	0.35	0.38	0.41	0.44	0.48	0.51	0.55	0.60
Steep, >7%	0.39	0.42	0.44	0.48	0.51	0.54	0.57	0.61
<i>Pasture/Range</i>								
Flat, 0 - 2%	0.25	0.28	0.3	0.34	0.37	0.41	0.46	0.53
Average, 2 - 7%	0.33	0.36	0.38	0.42	0.45	0.49	0.52	0.58
Steep, >7%	0.37	0.4	0.42	0.46	0.49	0.53	0.56	0.60
<i>Forrest/Woodlands</i>								
Flat, 0 - 2%	0.22	0.25	0.28	0.31	0.35	0.39	0.42	0.48
Average, 2 - 7%	0.31	0.34	0.36	0.4	0.43	0.47	0.50	0.56
Steep, >7%	0.35	0.39	0.41	0.45	0.48	0.52	0.54	0.58

It is evident from Fig. 4-1 that the $c_{(T)}$ -values obtained from this study are spread around those of Chow et al. (1988), but are generally lower in magnitude. The $c_{(T)}$ -values obtained from this study range from 0.084 to 0.786, while the values from Chow et al. are between 0.28 and 0.57 (for the recurrence interval range of 10- to 200-years). However, the scatter associated with the latter data set is not known and hence not shown, so it is conjectured that they are curves fitted to the high side of the original data.

Although the plot shown in Fig. 4-1 was initially performed purely for comparative purposes, its result was eventually used in the validation exercise performed in Section 4.1.5. It turned out that the extension of the calibrated coefficients to ungauged catchments proved very difficult (see Section 4.1.5), and hence it was decided to use the runoff coefficients of Chow et al. (1988: 498) as an approximation of the calibrated coefficients achieved in this study.

4.1.4. Hydrograph time base-length

The use of flood peak and volume pairs for calibration in this investigation, from the runhydrograph method of Hiemstra and Francis (1979), was thought to have the added advantage in that complete design flood hydrographs could be calculated from runoff coefficients calibrated on this dataset. From the flood database computed for the calibration exercise, hydrograph time base-lengths B for each of the return periods were determined from the peak-volume pairs for each catchment. These base-length values were then expressed as ratios to the catchment's time of concentration T_c for each of the respective recurrence intervals (which, in terms of the rational formula, is effectively a ratio to the hydrograph's time to peak). The average ratio of B/T_c , for each recurrence interval, was then determined and the results are presented in Table 4-3 together with their standard deviations. These results exclude three catchments whose area is 130 km² or less, as they gave B/T_c ratios in excess of 7. It is noted here that there is an increase of base-length with recurrence interval, which means that the volumes of the floods relative to the peaks, as modelled by the runhydrograph, also increase with T . The values in the third row of Table 4-3 show the proportion of floods whose base-length B exceeds $3T_c$, which is the value suggested by Rooseboom et al. (1981) and Alexander (2002), so that when T is 100, the proportion is approximately one third.

Table 4-3. The mean and standard deviations of the ratio of the hydrograph time base-length B to the catchments' time of concentration T_c as a function of recurrence interval T . The proportion of B/T_c values above 3 in each interval is given in the third row.

Recurrence Interval T (years)	10	20	50	100	200
Mean of B/T_c ratios	1.92	2.06	2.25	2.40	2.56
Standard deviation	0.981	1.09	1.29	1.48	1.71
Proportion > 3	0.14	0.19	0.28	0.34	0.40

4.1.5. Validation of the calibrated runoff coefficients

The purpose of validation is to test whether the model operates in the manner for which it was designed in "ways that were not explicitly built into the model" (Basson et al., 1994: 168). Validation tests are necessary to convey confidence that the model works as expected. In order to validate the $c_{(T)}$ -values achieved in calibration, it was necessary to find some physical or regional descriptor(s) on which to regress the coefficients. This was required so that the calibrated coefficients may be extended to ungauged catchments.

Several regional descriptors were tested in combination with the $c_{(T)}$ -values to examine if a relationship existed on which to regress the coefficients. Descriptors such as catchment slope, mean annual precipitation (MAP), percentages of land coverage and Kovačs' regional K -values (Kovačs, 1988) were tested. From these analyses, no meaningful relationships between any of the descriptors tested and the $c_{(T)}$ -coefficients were found. There were also no relationships found between parameters (multiplier and exponent) of a power-law function fitted to the $c_{(T)}$ -values, as a function of recurrence interval, and the regional descriptors. This result is in line with the comments of Pilgrim and Cordery (1992) for conditions in Australia, where the calibrated runoff coefficients did not show much sensitivity to catchment characteristics. However, this does indicate that the $c_{(T)}$ -values are essentially functions of T and T_c as conjectured. In light of the lack of dependency of the calibrated coefficients with catchment properties, an alternative solution was sought in order to extend the calibrated coefficients to ungauged catchments in validation.

Thus for the purposes of validation, it was decided to use the curves from Chow et al. (1988: 498), given in Table 4-2 and shown in Fig. 4-1. It can be seen from Fig. 4-1 that

the calibrated coefficients are generally lower than those of Chow et al. and, knowing that a practitioner will usually make a conservative choice, the latter coefficients were then viewed as an appropriate set of curves.

Twenty one catchments which were not used in calibration (and for which flood records were available) were selected for the validation exercise. These catchments ranged in size from 126 km² to 24 044km². From a previous study (see Pegram and Parak (2004) in Appendix A and Chapter 3), flood records were available and modelled (using a General Extreme Value (GEV) Distribution) for these catchments. The times of concentration (T_c) values were obtained from Petras and du Plessis (1987) and representative design rainfall intensities from Smithers and Schulze (2003) as before. These data are summarised in Table A7 (parts 1, 2 and 3) in Appendix B.

In order to obtain appropriate c-values from Chow et al. (1988: 498) for each catchment, it was necessary to relate the land coverage type and slope of each catchment with the descriptions of Chow et al. (see Table 4-2 above). These catchment characteristics (land coverage and slope) are given in Petras and du Plessis (1987) for each catchment. The land coverage types from this latter resource were catalogued as *forest, dense bush wood, thin bush wood, cultivated land, grass and bare*. At this stage it then became necessary to relate each catchment's coverage type to the generalized coverage descriptions of Chow et al. In order to easily accomplish this, two assumptions were made. They were: 1) that the greatest percentage of land coverage (the modal type) was representative of the entire catchment, and 2) the following coverage types (from the descriptions of Petras and du Plessis and Chow et al. respectively) were equivalent (shown in Table 4-4 below). It must be conceded that these assumptions had the tendency to be crude.

Table 4-4. Equivalent land coverage types from the descriptions of Petras and du Plessis (1987) and Chow et al. (1988: 498) for the catchments used in validation.

Equivalent land coverage types	
Actual catchment land coverage, as described in Petras and du Plessis (1987)	c-coefficient land coverages, as listed in Chow et al. (1988: 498)
Forest	Forest/Woodland
Dense Bush Wood	Forest/Woodland
Thin Bush Wood	Forest/Woodland
Cultivated Land	Cultivated Land
Grass	Pasture/Range
Bare	Cultivated Land

From the procedure described above, design flood peaks were obtained using the rational formula method (Q_{RF}) probabilistically, i.e. a function of catchment area, T -year design rainfall (of duration equal to the catchment time of concentration) and the corresponding T -year runoff coefficients from Table 4-2. These design flood peaks were then compared with the statistically modelled flood peaks (Q_{GEV}), from the same catchments, for the corresponding recurrence intervals. The results of this exercise, for the 10-, 50- and 200-year recurrence intervals are shown in Figs. 4-2, 4-3 and 4-4 respectively and are summarised for all recurrence intervals in Table 4-5.

Table 4-5. A summary of the *power-law* curves, of the form $Q_{RF} = aQ_{GEV}^b$, fitted to the graphs of Q_{RF} vs. Q_{GEV} of Figs. 4-2 to 4-4 and others (not shown). Q_{RF} are the flood peaks obtained from the probabilistically applied rational formula and Q_{GEV} are statistically modelled flood peaks. The *coefficient of determination* (R^2 -value) and the average ratio of Q_{RF}/Q_{GEV} for each recurrence interval are given in the last two rows respectively.

Recurrence Interval T (years)	10	20	50	100	200
Factor: a	5.44	5.10	5.17	5.75	7.03
Exponent: b	0.795	0.798	0.785	0.766	0.735
R^2	0.751	0.746	0.726	0.699	0.657
Mean Q_{RF}/Q_{GEV}	1.84	1.64	1.42	1.31	1.21

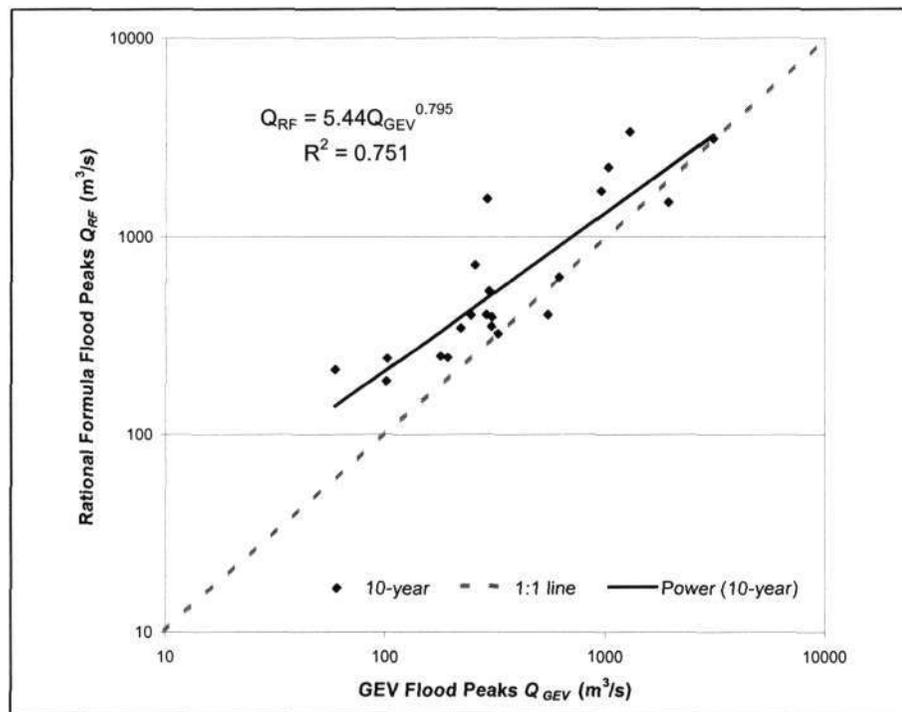


Figure 4-2. Plot in *log space* of the 10-year probabilistic rational formula flood peaks Q_{RF} versus the 10-year GEV modelled flood peaks Q_{GEV} .

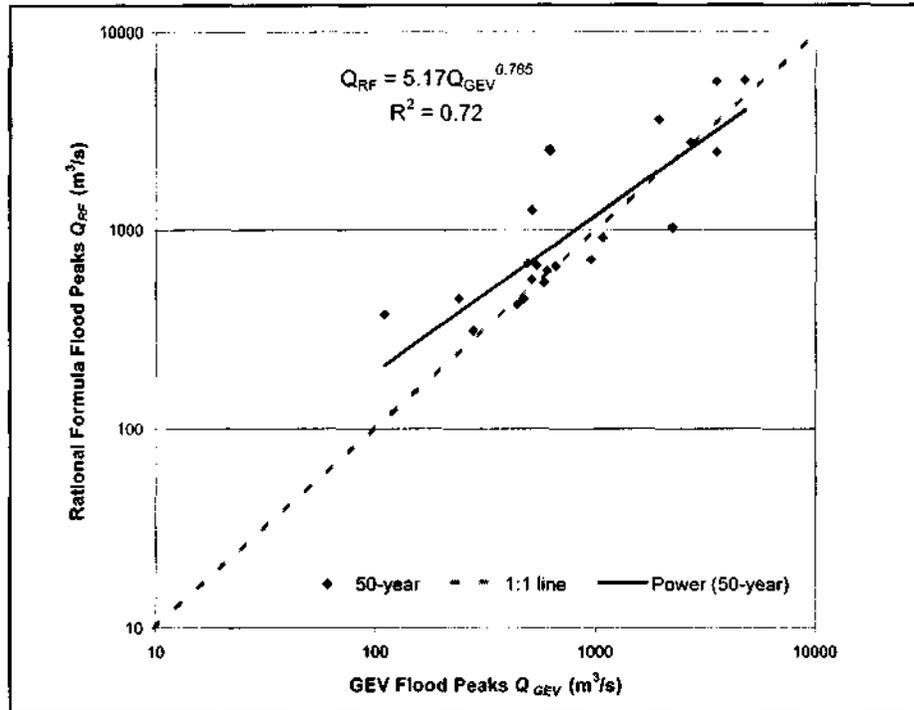


Figure 4-3. Plot in *log space* of the 50-year probabilistic rational formula flood peaks Q_{RF} versus the 50-year GEV modelled flood peaks Q_{GEV} .

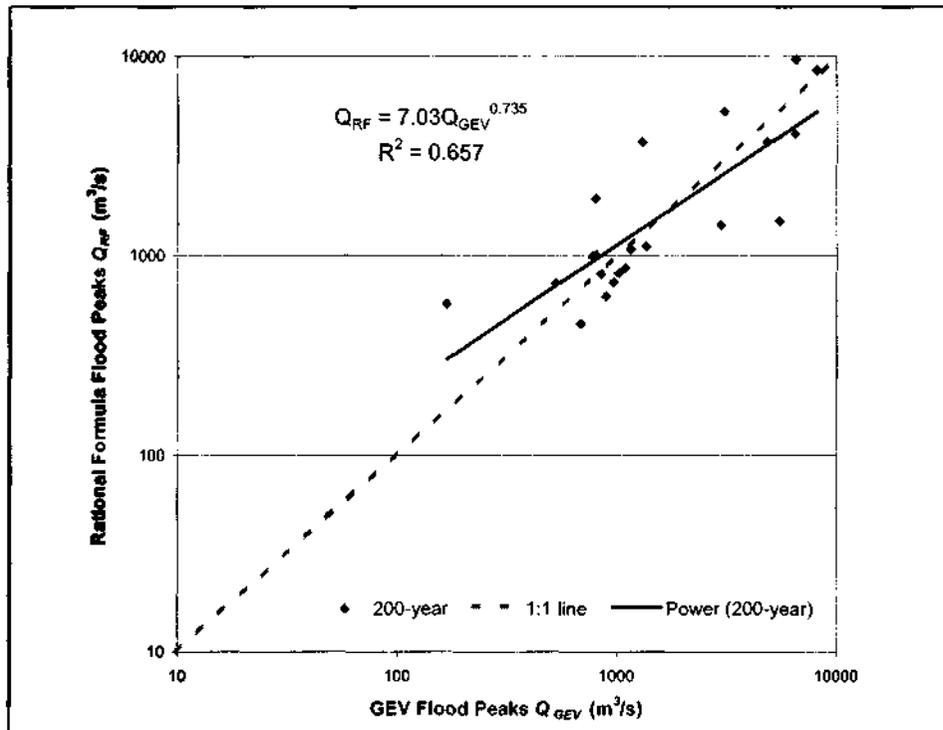


Figure 4-4. Plot in *log space* of the 200-year probabilistic rational formula flood peaks Q_{RF} versus the 200-year GEV modelled flood peaks Q_{GEV} .

Although there is a fairly large scatter around the trend-lines in log-space in Figs. 4-2, 4-3 and 4-4, some conclusions can be drawn from this validation exercise. It is evident

from these graphs that the estimated rational formula flood peaks Q_{RF} tend to be larger than the GEV modelled flood peaks Q_{GEV} , especially for the lower magnitude floods. However, at the larger flows, (approximately 7000 m³/s) the trend-lines cross the 1:1 line. This tendency is also exhibited for the 20- and 100-year validation tests (the results of which are not shown here) and is confirmed in Table 4-5 where the average ratio of Q_{RF}/Q_{GEV} across all recurrence intervals is approximately 1.5 (reducing from 1.84 for $T = 10$ to 1.21 for $T = 200$). This observation is to be expected since the c -values used to compute Q_{RF} in this exercise (from Chow et al. (1988: 498)), were generally larger than the calibrated runoff coefficients obtained in this study (see Table 4-2 and Fig. 4-1). Although the *coefficients of determination* (R^2 -values) are reasonable, the correlation is calculated in log-space and may disguise the fact that some flow peak ratios are occasionally different by up to a factor of 5 (see Table A7, Part 3 in Appendix B). As a consequence, the c -values adopted for this validation exercise, from Chow et al., were treated as upper bound estimates, conceding that although consistent, the method is prone to error.

4.2. Discussion of results

4.2.1. Calibration

Calibration of the rational formula's runoff coefficients, using runhydrograph flood peak and volume pairs of given recurrence intervals, was performed with the intention of removing some of the subjectivity involved in this parameter's estimation in the design environment. Use was made of T -year flood peak and volume pairs together with T -year design rainfall intensities, as a function of the catchments time of concentration, in order to obtain the coefficients. The results of this exercise produced calibrated runoff coefficients, as a function of recurrence interval, which were scattered (see Fig. 4-1) around published values from Chow et al. (1988: 498). The calibrated values, although spread around the latter set of coefficients, were in general lower in magnitude (with the exception of two catchments) and had gentler growths as a function of recurrence interval. Although this result did not produce a good match, the calibrated coefficients were sensible in magnitude. However, it was worrying to note that calibrated coefficients from six catchments (of the original 29 used in this exercise) had a tendency to decrease in magnitude with increasing recurrence interval. This deviation from the norm is attributed to the fact that the flood runoff data (calculated using the

runhydrograph method) had a gentler growth curve, as a function of recurrence interval, than the design rainfall data.

Although some effort at regionalisation was made, it was also found that the fitted c -values could not be regionalised and directly extended to ungauged catchments. This result was in agreement with the conclusions of Pilgrim and Cordery (1993). Since the fitted c -values (Fig. 4-1) were generally lower than those suggested by Chow et al. (1988: 498), it was therefore decided to accept the latter values as upper bound estimates for the purpose of validation, conscious of this discrepancy.

4.2.2. Hydrograph time base-length

It was initially thought that this investigation would be able to produce entire design hydrographs (albeit in an idealized triangular form) from the rational formula since the flood data used (from the runhydrograph method) described characteristic peak and volume pairs for each catchment. It was hoped that the ratio of B to T_c (effectively a ratio of B to the *time to peak* of a rational formula hydrograph) would be consistent and that a particular outflow hydrograph could be prescribed with the use of this method. However, the results (see Table 4-3) indicate that, firstly the average ratios are not constant across all recurrence intervals and, secondly that the coefficients of variation are quite high (they range from 0.51 to 0.66). Also, the results shown in Table 4-3 exclude three catchments of area less than 130km² as they gave ratios in excess of seven.

However, several points are worth noting from this exercise. Firstly the base-lengths are, on average, 2.25 times the catchments' time of concentration across all recurrence intervals. This result is somewhat less than the length of the hydrograph suggested by Rooseboom et al. (1981) and Alexander (2002), which was $3T_c$. However, Table 4-3 also indicates that a fair proportion of the calculated base lengths exceeded $3T_c$. As explained earlier, the hydrograph shape suggested by Rooseboom et al. (1981) was not meant to maintain continuity, but was instead designed to be conservative. The hydrographs derived in this study are thus expected to have a smaller base-length as continuity is implicitly maintained; so the result is in line with expectation.

Secondly, the tendency of the base-length to increase with T is possibly due to the method employed by Hiemstra and Francis (1979) in extracting their hydrographs and the non-linearity of the rainfall runoff process (rainfall losses reduce with increasing

recurrence interval). As depicted in Fig. 4-5, Hiemstra and Francis (1979) employed a truncation level for each catchment in order to extract independent hydrographs from their continuous records of stream flows. Flood volumes were obtained by extrapolating the rising limb and the recession limb of the discharge curves downwards towards zero flow from the first point below the truncation level which showed a reversal in slope. Depending on this level, a higher truncation level is likely to result in a reduction in the modelled volume when compared to the actual volume of the flood event. Thus it is likely that the base-lengths achieved in this study are smaller (as a function of T_c) for the smaller floods (more frequent events) than the base-lengths for the larger events, thus exhibiting the trend in Table 4-3.

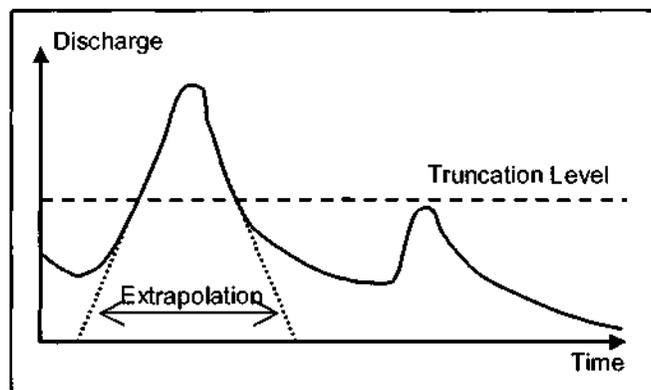


Figure 4-5. The method employed by Hiemstra and Francis (1979) to extract independent hydrographs from a continuous flow record, showing that a lower truncation level is likely to provide a bigger volume.

Finally, it is interesting to examine the relationship between B and T_c using a linear rainfall-runoff model as a comment on the values appearing in Table 4-3. If a constant (pulsed) input of rainfall of intensity i (in mm/h) on a catchment of area A (in km) lasts for the time of concentration T_c (hours), the total volume of rain that falls is $V = 1000 \cdot i \cdot T_c \cdot A$ (in m^3). The average rate of flow onto the catchment is $1000 \cdot i \cdot A$ (in m^3/h) and the peak outflow Q must be a fraction of this, say $\alpha \cdot 1000 \cdot i \cdot A$ (m^3/h), where $0 < \alpha < 1$. α is a factor related to the closeness of the peak to its asymptotic value as defined by its nearness to equilibrium. The base-length of the equivalent triangular hydrograph is thus $B = \frac{2V}{Q} = 2 \cdot T_c / \alpha$ (in hours). If there are no losses, the maximum peak that occurs at T_c can only be approaching equilibrium asymptotically, so α has to be chosen close to 1. If $\alpha = 0.9$, then it turns out that $B \approx 2.2T_c$, which is close to the average ratio (determined from Table 4-3 above), giving some support to the consistency of these results.

4.2.3. Validation

The validation exercise was necessary to test whether the calibrated coefficients behaved in the probabilistic manner for which they were designed, i.e. to predict design floods of magnitudes equivalent to those derived from a statistical analysis of flood records from that site. However, since it was shown that $c_{(T)}$ was not dependent on any physical properties nor region, c -values from Chow et al. (1988: 498), which are a function of return period, catchment slope and land-use characteristics, were substituted for the calibrated coefficients as approximate upper bound values. Based on this substitution, the validation exercise was ultimately reduced to a test of whether the c -values from Chow et al. (or possibly some other summary values) could provide reasonable design flood estimates such as those obtained from a statistical distribution (such as the GEV) fit to historical flood data.

The result of this exercise showed that the use of the substitute c -values from Chow et al. produced flood peaks from the rational formula that were, on average, approximately 1.5 times larger than floods estimated from a frequency analysis of historical data (see Table 4-5 and Fig. 4-2, 4-3 and 4-4), a result which is consistent with the c -values displayed in Fig. 4-1. The figures also show that the floods determined in this manner had a tendency to overestimate the statistically derived floods for lower flows and return period. This result is in line with expectation as the substitute c -values from Chow et al. were adopted as upper bound estimates. Given that, in order to make use of the coefficients of Chow et al., a crude matching of land coverage types was performed (see Table 4-4), this result is relatively pleasing especially since the catchments used in validation ranged in size from small to large (170 to 24000 km² – see Table A7 in Appendix B). However, the precision of the method is of course still low (as indicated by the spread of results in Figs. 4-2, 4-3 and 4-4) and still relies heavily on the judgement of the practitioner.

4.3. Chapter summary

Chapter 4 presented a review of the rational formula, by attempting to calibrate the most uncertain variable of the formula i.e. the runoff coefficient c . The results of the calibration were reasonably encouraging, producing c -coefficients that were scattered around, but generally lower than, those offered by Chow et al. (1988: 498), whose precision is not known. It was discovered that the fitted $c_{(T)}$ -coefficients of this

investigation did not show any variation with catchment characteristics, in line with Australian experience (Pilgrim and Cordery, 1993). Hence validation of these values at other sites was only possible through the substitution of c -values from Chow et al. as approximate upper bound estimates of the fitted $c(\eta)$ -coefficients. In order to use the values from Chow et al., a match of land coverage types was required. The results of the validation were as expected, producing floods from the rational formula that were on average 1.5 times larger than design floods estimated from a statistical analysis of historical streamflow records. However, it is noted that the results displayed wide scatter. Of lesser importance, it was discovered that the time base-lengths of the derived triangular hydrographs of this investigation were approximately between 1.9 and 2.6 times the catchment's time of concentration, depending on the recurrence interval of the flood. This result is lower than ratios suggested by Rooseboom et al. (1981) and Alexander (2002). It can be concluded from the results of this investigation that the probabilistic approach to the rational formula can be useful as a quick check method for calculating flood hydrographs for large catchments as it is for small.

CHAPTER 5

5. REAL-TIME FLOOD-FORECASTING USING RAINFALL-RUNOFF MODELS

This chapter is intended to briefly introduce the concepts associated with flood-forecasting, namely in the use of rainfall-runoff models in simulating catchment hydrology. The content presented in this chapter is not intended to be a comprehensive review on the subject matter, but as a basis for the introduction of a physically-based fully distributed rainfall-runoff model for real-time flood forecasting in Chapter 6.

As mentioned in the title of this chapter, this part of the research focuses on the application of a rainfall-runoff model for real-time flood-forecasting. The simulation of the rainfall-runoff process can be achieved through a variety of models which attempt to mimic catchment hydrology at various time and spatial scales. The scope of this research, with regard to these scales, is limited to catchments of area greater than 10km² and time's of concentration of 10 minutes and greater. The following sections introduce the concepts dealing with hydrologic modelling of catchments, its development and its application for this purpose.

5.1. Hydrologic modelling of catchments

A model is designed to be a tangible representation of a portion of the natural and/or anthropogenic world. In an engineering context, "a model is a set of physical laws written in mathematical terms and combined in such a way as to produce a set of outputs from a set of known inputs" (Haan, 1985). A hydrologic model is an approach which represents, mathematically, both the individual processes and all the interrelated processes involved in the hydrologic cycle (Martina, 2004). Hydrologic models of catchments are thus an assemblage of mathematical descriptions of components of the land-phase portion of the hydrologic cycle.

The components are generally modelled based on the conservation of mass, momentum and energy (physically based) or from a priori relationships (conceptual) of the phenomena. Some of the processes modelled in this manner are *inter alia* infiltration, evapotranspiration, groundwater percolation, unsaturated subsurface flow, saturated groundwater flow, overland surface runoff and channel flow. Fig. 5-1 depicts

typical processes of the hydrologic cycle. Some of the land-phase processes of the hydrologic cycle simulated in catchment models, such as infiltration, percolation, surface runoff and evapotranspiration, are also evident.

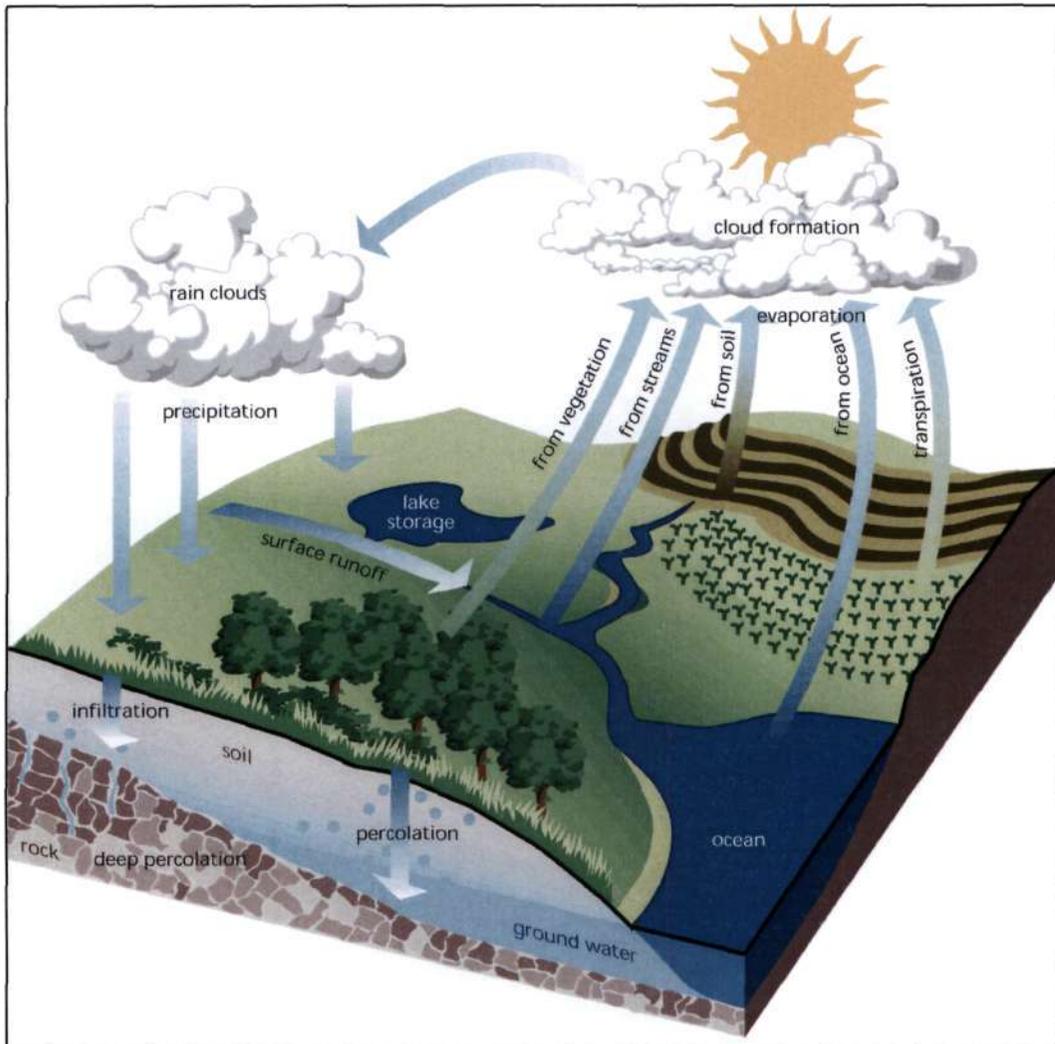


Figure 5-1. Typical processes of the hydrologic cycle (<http://www.cet.nau.edu/Projects/SWRA/research.html>).

5.1.1. Development of hydrologic models

The development of fully integrated catchment models only gained impetus during the middle of the 1960's, as explained by Singh and Woolhiser (2002). The 1960's saw the onset of the digital revolution, which made possible the integration of different components of the hydrologic cycle for the simulation of the entire catchment. Prior to this, hydrologic modelling involved the development of concepts, theories and models of the individual components of the hydrologic cycle. The first attempt to “model

virtually the entire hydrologic cycle” (Singh and Woolhiser, 2002) was made by Crawford and Linsley (1966) with the Stanford Watershed Model (SWM), which was a conceptual and lumped (semi-distributed) catchment model. The “blueprint” for a fully integrated, physically based three-dimensional model was proposed by Freeze and Harlan (1969) which included precipitation, surface runoff, porous media flow, open channel flow, the interaction of groundwater flow with channel flow and evapotranspiration. The 1970’s and 1980’s saw the introduction of remote sensing and geographical information systems (GIS) for data acquisition and management in catchment modelling. Singh and Woolhiser (2002) explain that little advancement was made through theoretical insights since the “blueprint” of Freeze and Harlan (1969). This model could not be implemented at the time because of computational and data limitations. Instead, many of the advancements made in this field since the 1970’s are due to the introduction of new measuring techniques and through the phenomenal development of computational facilities.

5.1.2. Concepts of hydrologic models

(a) Conceptual vs. physically-based models

Catchment models can be either classified in one of two ways, i.e. conceptual models or physically-based models. A model is said to be conceptual if it conceives of a system in which important hydrological processes are idealized (Schulze, 1995). Given reasonable a priori relationships, such as empirical representations, conceptual models simulate the physical reality in a simplified manner. By contrast, a model is deemed to be physically-based if the simulated phenomena of the hydrologic cycle are represented by governing equations which are deeply rooted in an understanding of the physics of the hydrologic cycle (Martina, 2004), for example through the conservation of mass, momentum and energy equations. The input variables of physically-based models should be physical quantities obtainable from direct observation or field measurement.

(b) Lumped vs. distributed modelling

Models can further be classified according to the spatial scale at which the processes are simulated, i.e. lumped or distributed models. A distributed model attempts to create a more faithful representation of reality by including the spatial distribution of the phenomena modelled. If the spatial detail is ignored, a model is said to be lumped.

Distributed models have the ability to predict the spatial trend of the hydrological conditions within a basin (Martina, 2004). By contrast lumped models are only able to produce outputs from inputs and no detailed insights into the internal distribution of the modelled processes are gained.

The idea of physically-based fully distributed models being a better approximation of reality than conceptual models was proposed by Freeze and Harlan (1969). However, Martina (2004) points out that there has been much debate on whether physically-based distributed models are feasible. These models require huge volumes of data and furthermore, as argued by Morel-Seytoux (1998), the nature of the phenomena modelled embodies both the elements of chance and the descriptive laws of physics. Therefore, excessive description at one scale is lost through the process of integration and averaging. However, a fully distributed physically-based model is necessary to gain a better understanding of the internal structure of the phenomena modelled. The challenge of physically-based distributed modelling is that the spatial and temporal considerations of modelling these processes need to be balanced with the computational burden of simulating the fully three-dimensional dynamics of catchment hydrology. Inevitably a few approximations of the governing laws, without affecting the physical meaning of the model, are necessary for mathematical tractability (Martina, 2004).

(c) Data requirements

The type of model to be built depends on the intended use of the model and the data available. Data requirements can be grouped into six categories and the data required in each category are listed in Table 5-1; the data are not specific to any particular model or model type. They are intended to show the general requirements associated with catchment modelling and any particular model may not require all of the data listed.

Table 5-1. General data requirements for catchment models (from Singh and Woolhiser, 2002).

Hydrometeorologic	Agricultural	Pedologic	Geologic	Geomorphologic	Hydrologic
Rainfall	Vegetative cover	Soil type	Stratigraphy	Topographic maps	Flow depth
Snowfall	Land use	Soil texture	Lithology	River networks	Streamflow discharge
Temperature	Land treatment	Soil structure	Structural controls	Drainage areas	Base flow
Radiation	Fertilizer application	Soil condition	Extent of aquifers	Slopes	Interflow
Humidity		Soil particle size		Slope lengths	Stream-aquifer interaction
Vapor pressure		Soil porosity		Catchment area	Potential water table
Sunshine hours		Soil moisture content			Drawdowns
Wind velocity		Capillary pressure			
Pan evaporation		Steady state infiltration			
		Saturated hydraulic conductivity			
		Antecedent moisture content			

(d) Data acquisition and manipulation

Distributed models require large quantities of data that need to be stored, retrieved managed and manipulated (Singh and Woolhiser, 2002). Through the use of remote sensing technology (radar and satellite) and geographical information systems (GIS), the ability to observe and map data over large spatial and temporal ranges is vastly improved. Together with the computing capability currently available, the development of fully distributed physically-based models is practically feasible.

(i) Remote sensing

Remote sensing uses measurements of the electromagnetic spectrum to characterize, infer properties and in some cases actually measure *inter alia* the magnitude and spatial distribution of meteorological inputs, soil and land-use parameters and the mapping of spatially varying landscape attributes. This technology goes a long way to alleviate the scarcity of data and provides the data in sufficient spatial detail, which was one of the greatest problems in the distributed hydrologic modelling of catchments (Singh and Woolhiser, 2002).

The use of remote sensing technology for real-time flood forecasting is important since meteorological inputs such as precipitation, or pedologic inputs such as soil moisture, vary temporally and spatially. Real-time applications of satellite and radar give an almost instantaneously available picture of where the precipitation is occurring in fine detail over a large area (Pegram and Sinclair, 2002). Examples of radar and satellite precipitation estimates are shown in Figs. 5-2 and 5-3. Fig. 5-2 shows a radar image of rain depth (mm) accumulated for the 24-hour period from 08:00 21/06/2005 to 08:00 22/06/2005. This image was taken from the South African Weather Services (SAWS) Website (<http://metsys.weathersa.co.za/simar-archive.html>) and superimposed on a digital elevation model (DEM) of Southern Africa (HYDRO1k, 1996). Fig. 5-3 shows a satellite image (from Pegram et al., 2005) of rain depth (mm) accumulated for the same 24-hour period of Fig. 5-2 taken from the Meteosat-8 Satellite. This image is also superimposed on the DEM as before. The rainfall resolutions shown in Figs. 5-2 and 5-3 are 1km and 1.6 km square (which is approximately 1 arc-minute at this latitude) for the radar and satellite images respectively while the resolution of the DEM is 1km square.

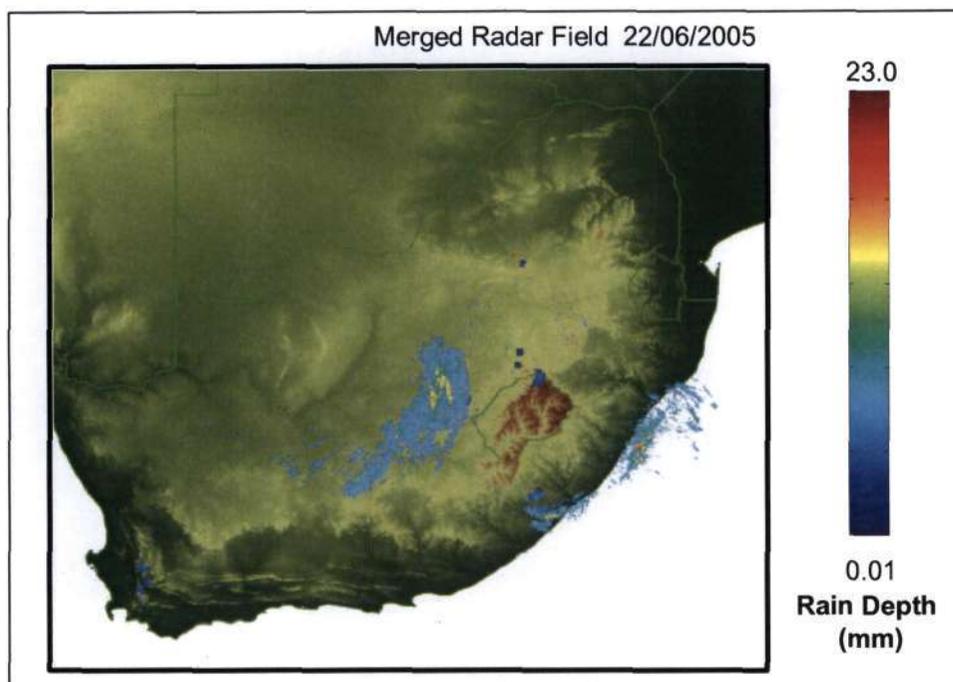


Figure 5-2. A radar image of rain depth (mm) accumulated for the 24-hour period from 08:00 21/06/2005 to 08:00 22/06/2005 (<http://metsys.weathersa.co.za/simar-archive.html>) superimposed on a digital elevation model (DEM) of Southern Africa (HYDRO1k, 1996).

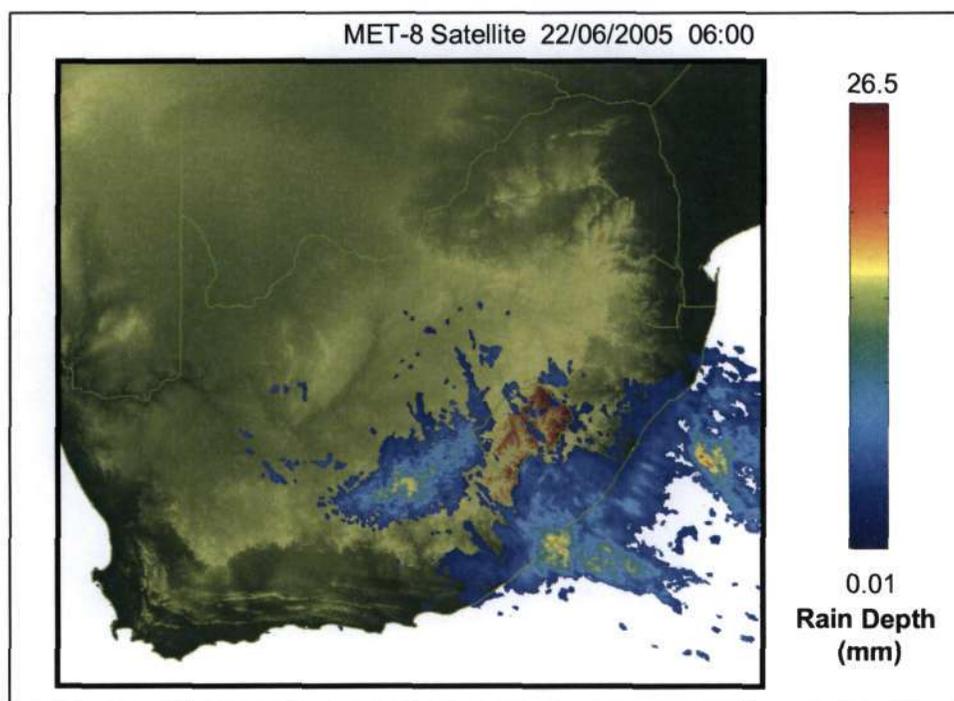


Figure 5-3. A satellite image of rain depth (mm) accumulated for the 24-hour period from 08:00 21/06/2005 to 08:00 22/06/2005 taken from the Meteosat-8 Satellite (Pegram et al., 2005) superimposed on a digital elevation model (DEM) of Southern Africa (HYDRO1k, 1996).

The spatial extent of the precipitation captured by the satellite and radar depicted in these Figs. 5-2 and 5-3 can be seen to be different. This difference highlights some of the errors associated with such measurement techniques, and no one technique can be taken to be the truth. In the example of Fig. 5-2, the altitude of the rainfall, especially at the coast, was too low to be captured by the radar's beam and secondly there is evidence of ground clutter (dark pixels) in the image. Thus it becomes necessary to combine the different precipitation estimates from remote sensing with block-Kriged telemetering raingauge data, which is taken as being the most representative of the truth, in order to obtain the "best" estimate of precipitation. This can be achieved using Bayesian combination (Mazzetti and Todini, 2002) or conditional merging (Pegram et al., 2005) and the product of this then becomes the input for real-time catchment models.

A further application of this technology is in the spatial mapping of landscape attributes. Different sensors of the electromagnetic spectrum are able to detect and/or measure (directly or indirectly), with a range of precision and scales, different properties of the landscape such as *inter alia* clouds, land-use, soil and vegetation characteristics, geology, temperature, soil moisture and water vapour. With reference to Table 5-1, many of the properties measured through remote sensing are essential, depending on the application, in the distributed modelling of catchments.

(ii) Geographical Information System (GIS)

A GIS is a computer-based system designed for retrieving, analyzing, manipulating, modelling, presenting and disseminating geographically-based data. It is an electronic system of maps connected to tables of data that describe features on maps (Dodson, 1992). A GIS enhances one's capability to incorporate spatial details in hydrologic models of catchments and has become an indispensable tool in the representation of the three-dimensional nature of landscape features in two-dimensions.

There are a number of ways of describing landscape features in a GIS. Features can be described discretely, using points, lines and polygons (vector-based GIS), where the coordinates, spatial relationships and characteristics of the features are recorded in relational tables. Vector data are well suited to recording geographic data that have precise locations (Chilufya, 2005). Alternatively, it is very useful for modelling applications to store information about geographic features that vary continuously over a surface, such as elevation data. Raster-based GIS is used for this and records spatial

information in a regular grid or matrix as a set of rows and columns (Chilufya, 2005). The size and shape of the grid cells are uniform, to which numerous landscape attributes can be attached, and the geographical location of each cell is known (Dodson, 1992). Since the attributes within each cell are assumed homogeneous, the accuracy of raster data is highly dependent on the resolution of the grid cell. The pixel (grid cell) format of digital remote sensing data makes raster-based GIS an ideal platform from which to manage and analyse this source of data (Dodson, 1992 and Engman, 1992).

Geomorphological information is an important ingredient in the accurate modelling of catchment hydrology (refer to Table 5-1). Prior to 1980, the main source of this information was contour mapping. Today, digital elevation models (DEMs) are used which are able to automatically extract the topographic variables required, such as basin geometry, stream networks, slope, aspect, flow direction, etc. (Singh and Woolhiser, 2002). A DEM is a raster-based representation of elevation data, i.e. it is a two-dimensional array of heights sampled above some datum that describe a surface. In many catchment models, as well as being a main source of topographic information, the pixel or grid cell of a DEM forms the primary processing unit for hydrological computations. Figs. 5-2 and 5-3 show, as the green background, a DEM of Southern Africa produced by the United States Geological Survey (USGS) at a resolution of 1km square (HYDRO1k, 1996), over which the radar and satellite precipitation images were superimposed using ARCGIS™ (a GIS software platform).

(e) Model scale

A successful catchment model must be able to reasonable duplicate two major hydrological processes occurring on the catchment, namely the conversion of rainfall to runoff and channel routing (Pegram and Sinclair, 2002). Many of the important phenomena simulated in physically-based distributed models, such as the process of infiltration for the conversion of rainfall into runoff, are described by means of balance equations in terms of the conservation of mass, momentum and energy at a point and an instant. In order to ensure compatibility between the observed input data and the governing equations, these point-scale equations need to be integrated to a finite scale dimension, since in general distributed catchment data are only available at finite dimensions. This upscaling, via integration, is an important link in converting the differential equations defined at a point into integral equations defined over volumes.

An important consideration at this juncture is the size of the grid dimension within which the physical meaning of the governing equations are still valid. Many models use the pixel size of a DEM as the primary processing unit. However, there has been much debate over whether the integration process preserves the physical-nature required in such models. As the size of the cells increases, the representative cell parameters tend to lose their physical meaning and become stochastic in nature. Wood et al. (1988) investigated the existence of a Representative Elementary Area (REA) which they defined as the minimum area within which "implicit continuum assumptions can be used without knowledge of the patterns of the parameter values, although some knowledge of the underlying distributions may still be necessary". This means that within each element, the variability of the parameters (parameter distribution) can be ignored and the element's characteristics can be considered relatively homogeneous. An assumption made at this scale is that the physical processes can still be represented in a deterministic way, as opposed to stochastic. Wood et al. (1988) found the REA to be strongly influenced by topography, but its upper limit approximately equals 1km^2 . Predictions from areas greater than this size were insensitive to the variability of parameters and therefore an REA of 1km^2 can be considered as the primary spatial unit.

Martina (2004) investigated the effects of integrating over the grid cell the process equations defined at a point for a fully distributed physically-based catchment model, i.e. the TOPKAPI model, which is introduced in Chapter 6. There were two issues that he looked at regarding this, namely the validity of the numeric method adopted to perform the integration and, secondly, the question of whether average parameters at the grid scale (which is necessary to ensure compatibility between the input parameters and the governing equations) are able to maintain physical meaning. In the first instance, it was found that the numerical solution of the flow equations adopted by the TOPKAPI model for the integration was valid up to an order of 1km^2 when compared to the "correct" solution offered by the *characteristic line method*. In the second instance, it was also found that the spatial distribution of the parameters within the grid cell can be averaged up to the order of 1km^2 . The result is in line with the findings of Wood et al. (1988) and suggests that, for the TOPKAPI model, "the approximations involved by the integration of the governing equations and by the consequent parameters averaging are acceptable" (Martina, 2004) at a grid resolution of up to 1km^2 . This REA size will be adopted in the sequel.

(f) Model computations

Many of the important processes in catchment modelling, such as the process of infiltration, are non-linear in nature. The governing non-linear differential equations at the finite grid scale may, in many instances, require a numerical or a pseudo-analytical solution. In a physically-based fully distributed model, such a solution only adds to the computational burden inherent in many models. A fundamental requirement of an online real time flood forecasting model is that it must be computationally efficient. Such a requirement has put the "feasibility" (Martina, 2004) of many physically-based distributed models in doubt and accounts for the popularity of model simplification (i.e. conceptual models). However, with the phenomenal computing power available today, the acquisition of suitable solutions is deemed not an obstacle to physically-based distributed models which are believed to represent reality more truthfully.

5.2. Chapter summary

In order to exploit the type and quality of data currently available, models for real-time flood-forecasting need to be physically-based and distributed. The computing capabilities also currently available make these types of models computationally sensible and efficient for this purpose. This chapter has introduced the concepts applicable to the subject matter and is the forerunner to the next chapter, where the TOPKAPI model, a fully distributed physically-based model is introduced and explained.

For the purposes of this research, the TOPKAPI model was chosen as the model to be used for the representation of the hydrologic phenomena of a catchment. The ultimate aim of the application of this model is for real-time flood-forecasting. However, being a fully distributed physically-based model, the application of such a model is able to represent the spatial trend and provide insight into the internal structure of the hydrologic phenomena of a catchment, for example for the calculation of soil moisture. Thus, such a model has a wide range of fields to which it is applicable. The TOPKAPI model was chosen for this purpose as it was considered to be superior to other models based on the fact that:

- it is a feasible fully distributed model, where the differential equations describing the physical processes of subsurface flow, overland flow and channel flow,

have been carefully integrated up and shown to be valid up to a grid resolution of 1km square, and

- it is also physically-based, where the input parameters, of which there are relatively few (seven), can (as claimed by its authors, Liu and Todini, 2002) be directly obtained from remote sensing and geographical information systems (GIS).

CHAPTER 6

6. TOPKAPI MODEL

Chapter 5 introduced the concepts associated with the real-time application of a rainfall-runoff model for flood forecasting. The need for truly physically-based fully distributed models was demonstrated in order to exploit the data and the computing facilities currently available. Several key concepts with regard to such models (fully distributed physically-based) were also discussed, namely the scale at which the physical nature of the phenomenon modelled is preserved, the computational burden of such models and the role that geographical information systems (GIS) and remote sensing play with regard to the supply and management of data required by the models.

To this end, the TOPKAPI model, a fully distributed physically-based hydrologic catchment model, was chosen for local application in a well instrumented catchment, i.e. the Liebenbergsvlei catchment. Since this model is novel to South Africa, this chapter provides an in-depth description of the model, its operations and its data requirements. The content contained in Chapter 6 is the result of an intensive dissection of disparate sources of literature on the model and the systematic combination of this information into a coherent whole. Through this literature study, various issues related to the model and its application are identified. These issues, such as the proportioning of channel flow, the local estimation of evapotranspiration and the solution of the non-linear differential equations describing the three stores of the model (soil, overland and channel store respectively) were either not clearly explained in the literature, or not suited to local conditions. These issues were either resolved or left for a follow-up study with recommendations, where it was deemed to be beyond the scope of this research. The contribution of Chapter 6 towards the aim of this research is the laying of a clear theoretical basis from which the model can be easily applied. Chapter 7 deals with the test application of the model. The primary sources of information, which were used to gain an understanding of the TOPKAPI model, were: Liu and Todini (2002), Bartholmes and Todini (2003), Martina (2004) and MUSIC Final Report (2004).

6.1. Description of the TOPKAPI model

TOPKAPI is an acronym which stands for **TOP**ographic **KIN**ematic **AP**proximation and **I**ntegration and is a physically-based distributed rainfall-runoff model. The model consists of five modules; a soil module, an overland module, a channel module, an evapotranspiration module and a snow melt module. A groundwater module is not incorporated in the current set-up of the model. The reason presented for this, as explained in Liu and Todini (2002), is that the response time for the percolation of water through the thick layer of soil separating the saturated (groundwater) and unsaturated (soil water) zones is relatively long. Hence the flow in the saturated zone shows no significant response from one storm event to another and can be assumed to be almost constant in time. However, the developers of the model plan to add a groundwater module for future enhancement of the model so that it can simulate all the hydrologic processes of a catchment (Liu and Todini, 2002).

The TOPKAPI model functions on the three main modules of soil, overland and channel in each processing cell, which is taken as the pixel of a digital elevation model (DEM). Each of these modules act as storages within a cell, receiving input and discharging output; the behaviour of the storages are described by structurally similar non-linear differential equations. The soil store is the regulatory storage of each cell, where its water balance determines the activation of the overland store. All precipitation input to a cell is infiltrated into the soil store. The overland store is activated upon the saturation of the soil store and both outflows from the soil and overland stores of a cell contribute a proportionate amount to the channel store of that cell. The remainder of the soil and overland outflows infiltrates into the soil store of the downstream cell while the channel outflow feeds the channel store of the downstream cell directly. Evapotranspiration is subtracted as a loss in each time step from the soil store of a cell. The non-linear differential equations describing each of the soil, overland and channel storages are derived by combining the *momentum* and *continuity* equations which describe the flow and storage of water. The non-linear storage equations are formulated by simplifying the momentum equations describing soil water flow (Darcy Law – see Section 6.1.2) and overland and channel flow (Saint-Venant Equation – see Section 6.1.3) through the application of the *kinematic wave model*. Darcy's Law describes flow in an unsaturated porous medium and the kinematic wave model assumes that the phreatic surface is equal to the topographic surface and that the effects of suction pressure are negligible in relation to gravity. The Saint-Venant Equation describes one-dimensional unsteady open channel flow and the kinematic

wave model assumes that all other forces acting on a control volume are negligible when compared to the effects of gravity and friction alone. These points are explained in detail in Sections 6.1.2 and 6.1.3. The two further modules, evapotranspiration and snow melt, are computed in the original TOPKAPI model as functions of temperature, the former using the radiation method of Doorenbos and Pruitt (1977), while the latter is also driven by a radiation estimate based upon air temperature measurements. The snow melt component of the TOPKAPI model is ignored in this research.

Thus through the combination of the continuity equation and the simplified momentum equation describing the flow and storage of water at a point, the resulting differential equations are then *integrated* in space to the finite dimension of a cell. The result is the non-linear differential storage equations which have been shown to have similar analytical solutions (Liu and Todini, 2002). The input parameters are directly obtainable from DEMs, soils maps and landuse maps for each cell of the catchment, in terms of slope, soil permeability, roughness and topology (Liu and Todini, 2002). The main advantage of the TOPKAPI model over other physically-based distributed models is that the physical nature of the governing equations and state variables are preserved in the integration process, albeit as averages, up to a grid scale of 1 km square (Martina, 2004: 76 and Martina et al., 2005). This satisfies the Representative Elementary Area (REA) criteria of Wood et al. (1988), referred to in Section 5.1.2 (e).

The objective of this research is in the real-time application of a rainfall-runoff model for flood-forecasting. The TOPKAPI model was chosen for this purpose as it was considered to be an appropriate model based on the following:

- it is fully distributed; the spatial range of the grid cell discretization within which the model is valid is up to 1km (Martina, 2004: 76),
- it is physically-based where the input parameters can be directly obtained from remote sensing and geographical information systems (GIS), and
- there are relatively few (seven) input parameters for a distributed model of which only three or four are typically used for calibration.

However, it must be noted that although the TOPKAPI model has been described as physically-based, and as such the input parameters should require no calibration in theory, Liu and Todini (2002) suggest that a calibration of parameters is still necessary. They maintain that the calibration of the model, which is “more an adjustment” achieved through simple trial and error methods, is still required because of the uncertainty of the information on topography, soil characteristics and land cover and

also because of the approximations introduced by the scale of the parameters representations. This is explained further in Section 6.1.7.

The major advantage in the application of a distributed model is its ability to represent the spatial behaviour of the phenomena modelled. However, in many applications, such as modelling the terrestrial - atmospheric flux, with Global Circulation Models (GCMs), models are required to represent phenomena at larger scales where the discretization schemes are based on grid sizes of hundreds of kilometres. The TOPKAPI model, although being a comprehensive distributed rainfall-runoff model, can also be applied in a lumped form to represent hydrologic processes at a basin level. This is achieved by using the distributed model to identify the mechanisms governing the dominant processes in the conversion of rainfall into runoff in order to obtain a "law underpinning the development of the lumped model" (Liu and Todini, 2002). It is shown in Martina (2004: 96) that the physical nature of the model is still maintained at the lumped scale, although the governing equations no longer have local meaning, but summarise local properties in a global manner where the input parameters represent basin averages. The application of the TOPKAPI model in lumped form is an option which is not explored in this research. The following sections explain the structure and methodology of the distributed TOPKAPI model by giving relevant background theory, expanded from the parent publications, where necessary.

6.1.1. Model assumptions

The TOPKAPI model is based on five fundamental assumptions. The reasoning behind each of the assumptions will become clearer as the explanation of the model is expanded in the sections that follow. Each assumption (appearing in italics) is quoted directly from Liu and Todini (2002) which is then followed by an explanation:

- (1) *Precipitation is constant over the integration domain, i.e. namely the single grid cell or pixel.* Precipitation estimates from remote sensing measurements are raster-based and presented in pixel format, where the properties, such as rainfall, are uniform in each pixel during the temporal integration interval. If raingauge data are used, suitable averaging techniques of these point estimates have to be performed in order to acquire the data at the required finite pixel resolution. One such technique that could be used for this purpose is Block Kriging.

- (2) *All precipitation falling on the soil infiltrates into it, unless the soil is already saturated in which case the input precipitation will become overland flow directly.* This is referred to as a Dunne mechanism (Dunne, 1978), or saturation excess mechanism, for the formation of overland flow.
- (3) *The slope of the groundwater table coincides with the slope of the ground, unless the latter slope is very small (i.e. < 0.01%).* This constitutes the fundamental assumption of the kinematic wave model, and justifies the use of the kinematic model for unsaturated horizontal subsurface flow.
- (4) *Local transmissivity, like horizontal subsurface flow in a cell, depends on the total water content of the soil.* This requires the integration of the soil water content profile in the vertical.
- (5) *Saturated hydraulic conductivity is constant with depth in the surface soil layer.* The conductivity in this layer is also assumed to be much larger than the conductivity in the deeper subsoil layers due to the macro-porosity that exists in this layer.

6.1.2. Soil water flow model

The soil water store is regarded as the most “characterising aspect of the model” (Liu and Todini, 2002) because of the regulating function that it plays in terms of the water balance. The overland flow component of a cell is activated upon the saturation of the soil store and together with subsurface flow, both contribute directly to the flow in the channel (Liu and Todini, 2002). Overland flow refers to surface runoff which flows down flat slopes in shallow sheets while subsurface flow, in terms of the TOPKAPI model, is regarded as the “flow in a horizontal direction that occurs in a soil layer of limited thickness and high hydraulic conductivity due to macro-porosity” (Liu and Todini, 2002). The TOPKAPI model assumes that flow in the vertical direction (typically over a depth of a few metres) of a given soil store, i.e. infiltration, is lumped and that the horizontal subsurface flow is a function of the total moisture stored within a soil store, which in turn is an explicitly derived function of the soil parameters. The following sub-sections deal with the kinematic wave formulation for the soil water flow model, by first introducing the necessary background theory, and then an explanation of the vertical lumping performed on this component.

(a) Background

The hydraulic behaviour of soil is characterised by two important properties of the soil, namely the *suction pressure head* ψ , which is the electrostatic force between the water molecules' polar bonds and the soil particle surface (Chow et al., 1988: 102), and the *hydraulic conductivity* K , which is the rate at which water moves through a porous medium per unit cross-sectional area. For unsaturated conditions, both these soil properties vary as a function of the *moisture content* θ of the soil. The soil moisture content θ is defined as the ratio of the volume of water to the total control volume (Chow et al., 1988: 100), with an upper bound limited by the porosity of the soil control volume. Porosity is defined as the ratio of the volume of voids (or pore spaces) to the total volume of the control volume. Therefore, at saturation, all the pore spaces within the control volume are occupied by moisture and hence the soil moisture content at saturation is equal to, or limited to, the porosity of the soil control volume. Fig. 6-1 shows the relationships of the suction pressure head and the hydraulic conductivity with moisture content for an unsaturated clay soil.

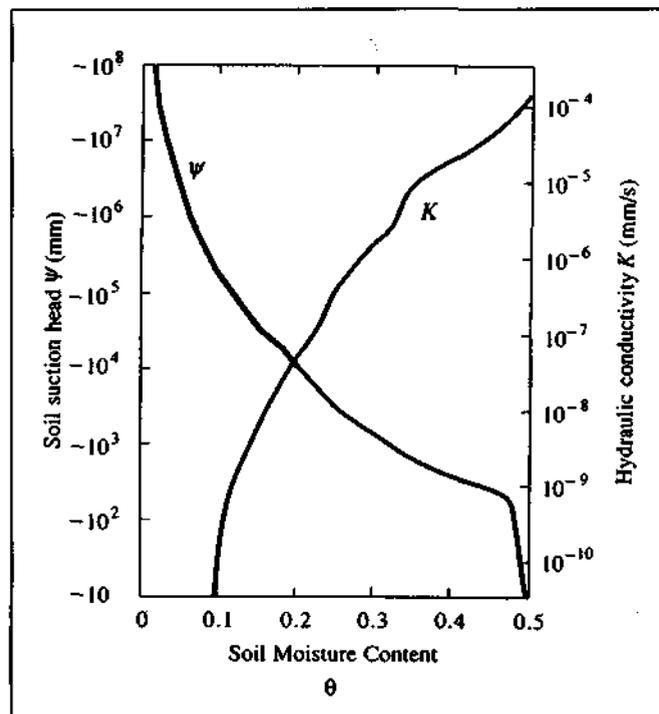


Figure 6-1. Variation of suction pressure head ψ and hydraulic conductivity K with moisture content θ for an unsaturated clay soil (Chow et al., 1988:103).

Flow in an unsaturated porous medium, such as subsurface flow, is described by Darcy's Law which gives the momentum equation for the volumetric flow per unit area of medium:

$$q' = KS_f \quad (6-1)$$

where q' is the apparent velocity of flow or volumetric flux (flow per unit area) through a cross section of a porous medium, K is the hydraulic conductivity of the cross section and S_f is the friction slope or head loss per unit length of flow in the x -direction. The friction slope is defined as $S_f = -\frac{dh}{dx}$, where the negative sign indicates that the head is decreasing in the direction of flow due to friction. For flow in an unsaturated medium, the forces involved are gravity, friction and suction head and hence the total head driving the flow are the sum of gravity and suction, i.e. $h = z + \psi$. (For saturated flow, the suction head is no longer applicable). Substituting for S_f in Eq. 6-1 for unsaturated flow yields:

$$q' = -K \left(\frac{\partial \psi + \partial z}{\partial x} \right) \quad (6-2)$$

where $(d\psi/dx)$ is the pressure head loss and (dz/dx) is the gravity head loss in the x direction respectively.

By combining the Darcy Equation, for flow in the vertical direction (z), with the continuity equation, Richard's Equation (1931) is derived, which is the governing equation for unsteady unsaturated flow in a porous medium. Richard's Equation is the basic theoretical equation for vertical flow in a porous medium, and is given as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left(D \frac{\partial \theta}{\partial z} + K \right) \quad (6-3)$$

where θ is the soil moisture content, t represents time, z represents the vertical direction (positive upwards), K is the hydraulic conductivity of the medium and D is a term which represents soil water diffusivity and equals $K(d\psi/d\theta)$. No general analytical solution to Richard's Equation exists due to its non-linearity. Infiltration models of Horton (1940) and Philip (1957) are approximate solutions to the one-dimensional form

of Richard's Equation. Green and Ampt (1911) proposed an alternative method whereby approximations were made with regard to the physical theory governing infiltration but yielding an exact solution (Chow et al., 1988: 110).

(b) Vertical lumping

In the TOPKAPI model, the dominant mechanism driving subsurface flow is assumed to be gravity and is described as follows, paraphrasing Martina (2004). Water, after having infiltrated into the soil, will perch on a lower impermeable or semi-permeable boundary of the soil layer. This boundary forms the separation between the subsurface soil layer and the deeper groundwater layer. At this boundary, a horizontal propagation of unsaturated subsurface flow is driven under gravity due to the relatively highly conductive (due to macro-porosity) nature of the soil layer. Since the depth of the soil layer (of thickness one to two meters) is negligible with regard to the horizontal dimension of the overall grid cell, according to Todini (1995) "it is possible to avoid within the range of reasonable errors the integration of the unsaturated soil vertical infiltration equation, namely Richard's Equation" (shown above as Eq. 6-3). Thus the TOPKAPI model assumes subsurface horizontal flow to be similar to flow in an unconfined aquifer, where "flows in unconfined aquifers are analogous to free-surface flows in streams" (Dingman, 2002: 327).

In order to lump the subsurface flow vertically, the dependence of the hydraulic conductivity K on the vertical soil moisture content $\theta(z)$ is neglected, which is noted here as the basic assumption of the Green-Ampt model (Green and Ampt, 1911). Rather, a direct dependency of hydraulic conductivity with the averaged soil moisture content over the soil depth is assumed to exist. This assumption was based on the fact that in practice, as explained in Martina (2004: 81), "the horizontal flux evaluated from the integration of the vertical profile of the soil moisture content does not differ strongly from the horizontal flux evaluated from assuming the saturated hydraulic conductivity as constant with depth", together with the average value along the vertical profile of the soil moisture content.

The formulation of the relevant equations is as follows. Firstly, the TOPKAPI model applies the kinematic wave approximation to the subsurface flow and thus the effects of suction on Darcy's Law are ignored in relation to gravity. Secondly, the model assumes that the slope of the phreatic surface (and thus the slope of flow perched on the lower impermeable or semi-permeable boundary) is equal to the slope of the topographic

surface β . Thus dz/dx in Eq. 6-2 equals $\tan\beta$, where β is the local angle of slope of the topographic surface of a cell in the direction of flow. Eqs. 6-1 and 6-2 thus reduce to:

$$q' = K \cdot \tan \beta \quad (6-4)$$

where K is the hydraulic conductivity of the soil and $\tan\beta$ is the tangent of the ground slope angle β . When dealing with saturated horizontal flow paths, such as in an unconfined aquifer, the term *transmissivity* is used instead of hydraulic conductivity. Transmissivity is defined as $T = H \cdot K$, where H is the saturated flow depth and K is the hydraulic conductivity. Thus K in Eq. 6-4 is replaced by T/H , and thus $q = H \cdot q' = H \cdot K \cdot \tan \beta = T \cdot \tan \beta$. Based on assumption (4), in Section 6.1.1, local transmissivity of the soil profile depends on the total moisture content of the soil over the depth L . This is calculated as follows:

$$T = \int_0^L K(\bar{\theta}(z)) dz \quad \left(= \int_0^L K_s (\bar{\theta}(z))^\alpha dz \right) \quad (6-5)$$

where K is the hydraulic conductivity as a function of the *effective soil saturation* $\bar{\theta}$ at depth z and L is the depth of the soil layer. The effective soil saturation at depth z is the ratio of the available moisture ($\theta - \theta_r$) to the maximum possible moisture content ($\theta_s - \theta_r$), i.e. $\bar{\theta}(z) = \frac{\theta - \theta_r}{\theta_s - \theta_r}$ (Chow et al., 1988: 114). θ_r is the residual soil moisture content after the soil has been thoroughly drained and θ_s is the saturated soil moisture content. The replacement of the integrand on the left of Eq. 6-5 with the approximation shown in brackets on the right is based on the work of Brooks and Corey (1964) who established the relationship $K(\bar{\theta}) = K_s (\bar{\theta})^\alpha$, where K_s is the saturated hydraulic conductivity and α

is a pore-size distribution parameter which is dependent on the characteristics of the soil type. They established this formula after studying the relationships between suction pressure head ψ and hydraulic conductivity K with the moisture content θ of many soils.

By assuming a constant saturated hydraulic conductivity K_s with soil depth and by using the average soil moisture content over the depth, the expression for the transmissivity given by Eq. 6-5 is replaced by:

$$T(\bar{\Phi}) = K_s L \bar{\Phi}^\alpha \quad (6-6)$$

where $\bar{\Phi}$ is the average effective soil saturation content over the depth L , i.e.

$$\bar{\Phi} = \frac{1}{L} \int_0^L \bar{\theta}(z) dz. \quad (\text{At this stage it is important to introduce a system of dimensions; the SI system will be used}).$$

Thus from Eq. 6-4, the horizontal subsurface flow q (in m^2s^{-1}) is can be expressed by:

$$q = \tan(\beta) \cdot K_s L \bar{\Phi}^\alpha \quad [\text{m}^2\text{s}^{-1}] \quad (6-7)$$

(c) Kinematic wave formulation for subsurface flow

The momentum equation (Eq. 6-1) and the continuity equation (Eq. 6-3) are used to obtain the following pair of equations that describe the flow and storage of soil moisture, lumped vertically in a column of depth L and of elemental area dA in the horizontal plane:

$$q = \tan(\beta) \cdot k_s L \bar{\Phi}^\alpha \quad [\text{m}^2\text{s}^{-1}] \quad (6-8a)$$

$$(\theta_s - \theta_r) L \frac{d\bar{\Phi}}{dt} = p - \frac{dq}{dx} \quad [\text{ms}^{-1}] \quad (6-8b)$$

where x is the horizontal direction of flow in a cell, t is time, q is the subsurface flow (in m^2s^{-1}), p is the input precipitation intensity (in ms^{-1}) and the rest of the variables are defined as before. The term on the left-hand side of the continuity equation, Eq. 6-8b, represents the rate of change of moisture storage (expressed as depth) in the soil store while the expressions on the right-hand side are the inflow and outflow balance. The model is written in terms of total differential operators instead of partial differential operators, since the flow in the TOPKAPI model is assumed to be characterised by a preferential direction which is defined as the direction of maximum slope.

By combining Eq. 6-8a and Eq. 6-8b, the resulting equation states that the rate of change of storage ($d\eta/dt$) is equal to the difference between the inflow (p) and outflow ($d(C\eta^\alpha)/dx$):

$$\frac{d\eta}{dt} = p - \frac{d}{dx}(C\eta^\alpha) \quad [\text{ms}^{-1}] \quad (6-9)$$

where η is the total depth (in m) of the actual moisture in the soil (defined by Eq. 6-10), C represents a local conductivity coefficient (defined by grouping the constant physical terms in Eq. 6-11) and the rest of the variables are as before. The expression for the soil moisture depth (Eq. 6-10) is based on the average effective soil moisture $\bar{\phi}$ of the soil layer of depth L :

$$\eta = (\theta_s - \theta_r) L \bar{\phi} \quad [\text{m}] \quad (6-10)$$

and

$$C = \frac{L \cdot k_s \cdot \tan \beta}{(\theta_s - \theta_r)^\alpha \cdot L^\alpha} \quad [\text{m}^{2-\alpha} \text{s}^{-1}] \quad (6-11)$$

Eq. 6-9 expresses continuity in 2-dimensions, i.e. for an elemental area over the horizontal plane of a grid cell but lumped in the vertical dimension. In order to represent the processes over an entire pixel or grid-cell, Eq. 6-9 needs to be integrated, firstly over the longitudinal dimension X and then over the width X of the grid-cell. Fig. 6-2 depicts the dimensions of a typical soil store of a grid-cell which is defined horizontally by the pixels of a digital elevation model (DEM).

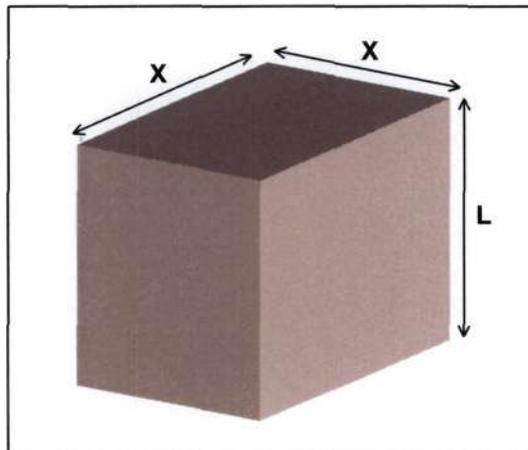


Figure 6-2. The dimensions (in m) of the soil store in a pixel as defined by a DEM and as required by TOPKAPI. The picture is not to scale and the dimensions have been exaggerated.

For the i^{th} pixel, assuming the pixel to contain a source-cell with no upstream cells contributing flow (i.e. precipitation is the only input into the cell), the integration of Eq. 6-9 in the longitudinal direction yields the following:

$$\int_0^X \frac{d\eta}{dt} dX = \int_0^X \left(p - \frac{d}{dX} (C\eta^\alpha) \right) dX \quad (6-12)$$

$$\Rightarrow \frac{dv_i}{dt} = pX - C_i \eta_i^\alpha \quad [m^2 s^{-1}]$$

where v_i is the volume per unit width ($m^3 m^{-1}$) stored in the i^{th} cell, i.e. $v_i = \eta_i \cdot X$ in m^2 . This assumes that the variation of the vertical water content η_i is negligible along the horizontal dimension of the cell and thus v_i can be related to the vertical water content η_i as shown above. After integration, the conductivity coefficient C no longer comprises measurable quantities at a point, but now represents average values over the cell. The total volume (in m^3) stored in the i^{th} cell is V_i , where $V_i = v_i \cdot X = \eta_i \cdot X \cdot X$. Thus, by making the substitution for η in Eq. 6-12 and integrating over the width of the i^{th} source cell, the non-linear reservoir equation for this store is:

$$\int_0^X \frac{dv_i}{dt} dX = \int_0^X (pX - C_i \eta_i^\alpha) dX \quad (6-13)$$

$$\Rightarrow \frac{dV_i}{dt} = p_i X^2 - \frac{C_i X}{X^{2\alpha}} V_i^\alpha \quad [m^3 s^{-1}]$$

where V_i is the volume stored in the i^{th} cell in m^3 . In a similar manner, a non-linear reservoir equation can be formulated for a generic "non-source" cell which, in addition to precipitation input, receives contributions from the soil and overland stores of the upstream cell.

The connectivity between cells in the TOPKAPI model is such that an active cell may only receive upstream contributions from the three cells adjacent to the edges of the active cell and may only have one "preferential" outflow direction. A further explanation of this is given in Section 6.1.5. Referring to Fig. 6-3, the non-linear reservoir equation for a generic cell is:

$$\frac{dV_{s_i}}{dt} = p_i X^2 + Q_o^u + Q_s^u - \frac{C_{s_i} X}{X^{2\alpha_{s_i}}} V_{s_i}^{\alpha_{s_i}} \quad [m^3 s^{-1}] \quad (6-14)$$

where V_{s_i} is the volume stored in the i^{th} cell and Q_o^u and Q_s^u are the direct contributions from the upstream overland and soil stores feeding the i^{th} cell respectively. These terms were effectively added to the right-hand side of Eq. 6-13 for a non-source cell. The subscripts o and s have been introduced to distinguish between the overland and

soil stores respectively and the superscript u denotes an upstream contribution. Note that Eq. 6-14 is a scalar non-linear differential equation in time.

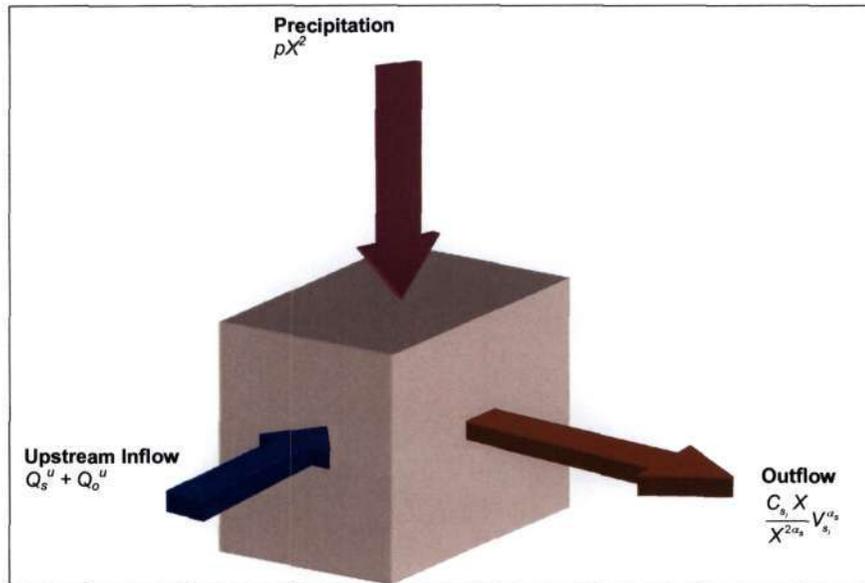


Figure 6-3. The water balance for the soil store of a generic non “source” cell derived in Eq. 6-14 for the TOPKAPI model (MUSIC Final Report, 2004: 53).

6.1.3. Overland and channel water models

The overland and channel components are controlled by the soil store. Overland storage is activated when the soil store is saturated, based on the saturation excess mechanism of Dunne (1978), and proportions of the outflow from the soil and overland stores feed the channel store. This section deals with the kinematic wave formulation for the overland and channel flow models by first introducing the necessary background open channel flow theory.

(a) Background

The kinematic wave assumption is based on the simplification of the momentum equation in the pair of Saint-Venant Equations which describe one-dimensional unsteady open channel flow. It assumes that the effects of local acceleration, convective acceleration and pressure acting on a control volume are negligible when compared to the effects of gravity and friction. The Saint-Venant Equation of *momentum conservation* is:

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} + g \frac{\partial y}{\partial x} - g(S_o - S_f) = 0 \quad (6-15)$$

where V represents the velocity in the x -direction at time t where x is the horizontal distance along the channel, g the gravitational acceleration and S_o and S_f the bed slope and friction slope of the channel respectively. This equation is most applicable in describing gradually varied unsteady flow and is typically used for bed slopes (S_o) flatter than 1:20 so that the horizontal component of the velocity is dominant. The kinematic wave approximation of the momentum equation ignores the effects of local acceleration ($\partial V/\partial t$), convective acceleration ($V \cdot \partial V/\partial x$) and pressure ($g \cdot \partial y/\partial x$). This simplification is justified by an observation of Henderson (1966: 364) where he notes that $\partial V/\partial t$, $V \cdot \partial V/\partial x$ and $g \cdot \partial y/\partial x$ are relatively small in comparison to S_o (0.2%, 1% and 2% respectively) for natural floods in actual rivers. Thus, by ignoring these terms in Eq. 6-15, the following approximation results:

$$S_f = S_o \quad (6-16)$$

where the forces of friction and gravity balance. By combining this approximation with the continuity equation at a channel cross section (Eq. 6-17), the flow and storage of channel water is described by a non-linear reservoir differential equation for each channel store in a processing cell. The *continuity equation* for a channel cross section, in conservation form is:

$$\frac{\partial A}{\partial t} = q - \frac{\partial Q}{\partial x} \quad (6-17)$$

where $(\partial A/\partial t)$ is the rate of change of volume per unit width in a channel element, q is the inflow per unit length along the side of the channel and $(\partial Q/\partial x)$ is the rate of change of channel flow with distance (Chow et al., 1988: 275). The formulation of the non-linear differential equation for the overland storage follows from the preceding arguments.

Thus, based on the kinematic approach for the conservation of mass and momentum, flow in the overland store (surface flow) and flow in the channel store (channel flow) are described by Manning's Equation. The Manning Equation is valid for one-dimensional steady, uniform open channel flows and is given as (in SI units):

$$Q = \frac{1}{n} \cdot A \cdot R^{2/3} \cdot S_f^{1/2} \quad (6-18)$$

where Q is the volumetric flow rate (m^3s^{-1}), n is Manning's roughness coefficient ($\text{m}^{-1/3}\text{s}$), A is the cross sectional area of flow (m^2), S_f is the friction slope (which is equal to the ground slope $\tan\beta$ in the kinematic approximation) and R is the hydraulic radius (in m), which is defined as the ratio of the cross-sectional area of flow (A) to the length of the cross-sectional wetted perimeter of flow (P), i.e. $R = \frac{A}{P}$. Referring to Fig. 6-4, the cross sectional area for a rectangular channel is yB and the wetted perimeter is $2y+B$. A common assumption made when using Manning's Equation is that the cross section of flow is rectangular with the width of flow B being much larger than the height of flow y , i.e. the channel is wide and rectangular. This assumption is valid when one considers overland flow as this flow is akin to sheet flow. However, when the flow is in a channel, the cross section of flow needs to be shallow and wide enough for the assumption to be valid. Nevertheless, Dingman (2002: 427) says that in most natural channels, the hydraulic radius R is "virtually identical to the average depth y ", which arises from the *wide rectangular channel* assumption. This assumption reduces Eq. 6-18 to the following approximation:

$$q = \frac{Q}{B} = \frac{1}{n} \cdot (\tan\beta)^{1/2} \cdot y^{5/3} \quad (6-19)$$

where q is the flow per width in the channel (in m^2s^{-1}).

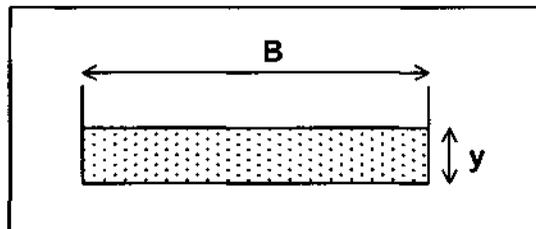


Figure 6-4. The rectangular cross section of an open channel flow of flow depth y and flow width B .

(b) Kinematic wave formulation for overland flow

The momentum equation (Eq. 6-19) and the continuity equation (Eq. 6-17) for the overland store are used to obtain the following pair of equations that describe the flow and storage of overland water at a point in the overland store:

$$q_o = C_o \cdot h_o^{\alpha_o} \quad [\text{m}^2\text{s}^{-1}] \quad (6-20a)$$

$$\frac{dh_o}{dt} = r_o - \frac{dq_o}{dx} \quad [\text{ms}^{-1}] \quad (6-20b)$$

where q_o is the overland flow per unit width, h_o is the depth of flow over the ground surface (in m), $C_o = \frac{1}{n_o} \sqrt{\tan \beta}$ and $\alpha_o = \frac{5}{3}$ (an overland flow parameter), both of which come directly from Eq. 6-19. In Eq. 6-20b, r_o is defined as saturation excess flow or runoff (in ms^{-1}) which results from the saturation of the soil store; evidently r_o is zero when the soil store is unsaturated. The subscript o denotes the overland flow model.

In a similar treatment to that of the soil store (the equations have the same form), Eqs. 6-20a and Eq. 6-20b can be combined and integrated over the horizontal dimensions of the grid cell by assuming that the depth of flow h_o is constant over the surface of the cell. Referring to Fig. 6-5, the non-linear reservoir equation for the overland store of a generic cell is given as:

$$\frac{dV_o}{dt} = r_o X^2 - \frac{C_o X}{X^{2\alpha_o}} V_o^{\alpha_o} \quad [\text{m}^3\text{s}^{-1}] \quad (6-21)$$

where V_o is the overland volume stored in the i^{th} cell, $\frac{dV_o}{dt}$ is the rate of change of surface water storage in the overland store, $r_o X^2$ is the input term and $\frac{C_o X}{X^{2\alpha_o}} V_o^{\alpha_o}$ is the outflow term.

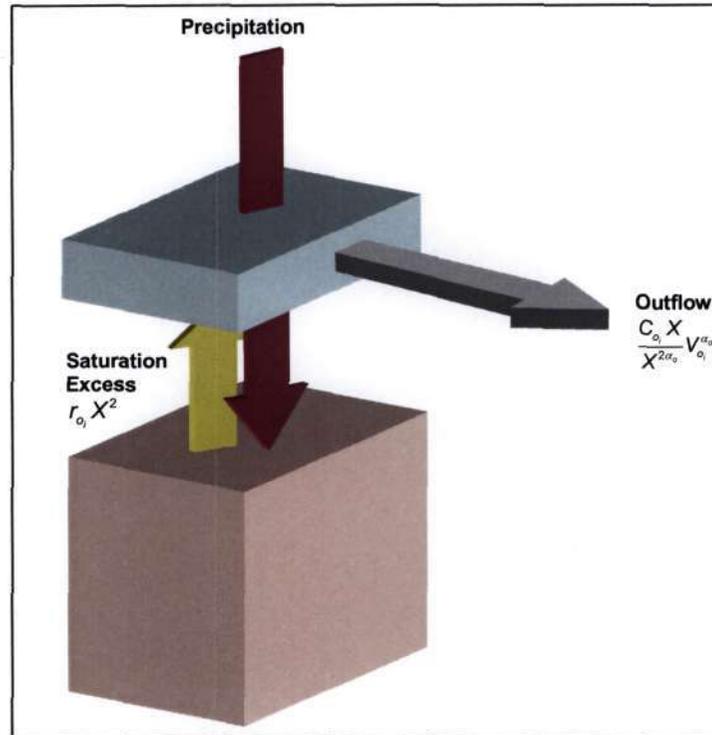


Figure 6-5. The water balance for the overland store of a generic cell derived in Eq. 6-21 for the TOPKAPI model (MUSIC Final Report, 2004: 55).

(c) Kinematic wave formulation for channel flow

The channel flow model is only applicable in those cells that contain a channel reach. This differentiation is determined from a topographic analysis of the catchment through the use of a digital elevation model (DEM). The distinction between cells characterised by overland runoff only (hillslope cells) and cells characterised by commensurate overland and channel runoff (channel cells) is important for the type of water balance to be created for each cell. This is dealt with in Section 6.1.5.

In a similar manner to the treatment of the soil and overland stores, the momentum equation (Eq. 6-19) and the continuity equation (Eq. 6-17) for the channel store are used to obtain the following system of equations that describe the flow and storage of water at a point in a channel reach:

$$q_c = C_c \cdot y_c^{\alpha_c} \quad [\text{m}^2\text{s}^{-1}] \quad (6-22a)$$

$$\frac{dy_c}{dt} = r_c - \frac{dq_c}{dx} \quad [\text{ms}^{-1}] \quad (6-22b)$$

where y_c is the depth of flow in the channel reach (in m), $C_c = \frac{1}{n_c} \sqrt{\tan \beta}$ and $\alpha_c = \frac{5}{3}$, both of which derives from Eq. 6-19 (again assuming that the cross section of flow in the channel reach to be wide and rectangular). In Eq. 6-22b, r_c is defined as the lateral drainage input (in ms^{-1}) which results from the contribution of the outflows of the soil and overland stores of the cell. The subscript c denotes the channel flow model.

Eqs. 6-22a and 6-22b can be combined and integrated over the horizontal dimensions of the channel reach in the grid cell by assuming that the depth of flow y_c is constant in the channel reach. The horizontal dimensions of a channel reach, as depicted in Fig. 6-6, do not occupy the entire width of the grid cell. The width W_i of a channel reach is assumed to remain constant over the entire length of the cell, but is larger in downstream cells increasing towards the channel outlet as a function of the area drained (see Eq. 6-24). Referring to Fig. 6-6, the non-linear reservoir equation for a channel reach in a generic cell is:

$$\frac{dV_{c_i}}{dt} = r_{c_i} X W_i + Q_c^u - \frac{C_c W_i}{(X W_i)^{\alpha_c}} V_{c_i}^{\alpha_c} \quad [\text{m}^3 \text{s}^{-1}] \quad (6-23)$$

where V_{c_i} is the channel volume stored in the channel reach of the i^{th} cell and Q_c^u is the channel inflow from an upstream cell. From Eq. 6-23, it is evident that the model requires that the proportion of soil and overland flow from the i^{th} cell that feeds its channel is proportionate to the ratio of the width W_i of the channel to the overall width X of the cell. If Q_{s_i} and Q_{o_i} are the soil and overland outflows available to feed the channel reach of the i^{th} cell in each time interval respectively, then $Q_{s_i} + Q_{o_i} = r_{c_i} \cdot X \cdot X$ and $r_{c_i} \cdot X \cdot W_i = (r_{c_i} \cdot X \cdot X) \cdot \frac{W_i}{X} = (Q_{s_i} + Q_{o_i}) \cdot \frac{W_i}{X}$ in $\text{m}^3 \text{s}^{-1}$. Thus the formulation of Eq. 6-23 guarantees that the amount of soil and overland flow available to feed the soil store of the downstream cell (cell $i+1$) is proportional to $\left(1 - \frac{W_i}{X}\right)$. This proportioning can be seen in Fig. 6-6. This point was not made clear in the literature studied and was interpreted here.

Not clear

The width W_i of the channel reach of the i^{th} cell is calculated as follows:

$$W_i = W_{max} + \left[\frac{W_{max} - W_{min}}{\sqrt{A_{total}} - \sqrt{A_{threshold}}} \right] \left(\sqrt{A_{drained_i}} - \sqrt{A_{total}} \right) \text{ [m]} \quad (6-24)$$

where A_{total} is the total area drained in the catchment, $A_{threshold}$ is the threshold area, which is the minimum drainage area required to initiate a channel, $A_{drained_i}$ is the area drained by the i^{th} cell, W_{max} is the maximum width of the channel (in meters) at the basin outlet and W_{min} is the minimum width corresponding to the threshold area. The parameters A_{total} , $A_{threshold}$ and $A_{drained_i}$ are determined from the topographic analysis of the DEM of the catchment while values for W_{max} and W_{min} are estimated based on a priori knowledge of the site or estimated from satellite imagery.

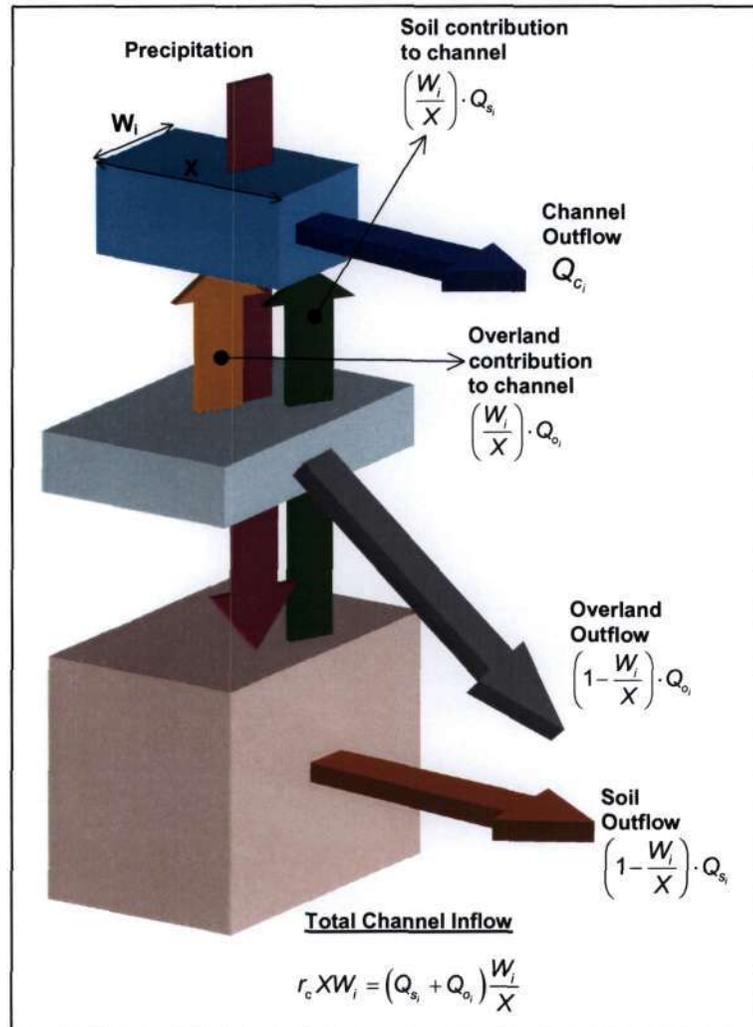


Figure 6-6. The water balance for the channel store of a “source” cell derived in Eq. 6-23 for the TOPKAPI model.

6.1.4. Evapotranspiration module

The evapotranspiration module in the TOPKAPI model is basically a moisture loss function and subtracts from the soil store an accumulated amount of moisture in each time step. The computation of evapotranspiration losses was not done dynamically in the integration of Eq. 6-14, the non-linear reservoir equation for a soil store, as its instantaneous impact was not considered important in the rainfall-runoff process. It was felt by the developers of the model, that evapotranspiration losses would have a small dynamic effect during a time step (of the order of less than an hour) and that it was only necessary to preserve the cumulative volumetric balance in order to maintain the correct soil moisture budget (MUSIC Final Report, 2004: 51).

(a) Background

Evapotranspiration is the combination of evaporation from the soil surface and transpiration from vegetation (Chow et al., 1988: 91). The factors that govern potential evapotranspiration are energy supply, vapour transport and the supply of moisture at the evaporative surface. The estimate of actual rate of evapotranspiration, for a given crop and climate, is based on the rate of evapotranspiration of a reference crop. The *reference crop evapotranspiration* (E_r) is defined as "the rate of evapotranspiration from an extensive surface of 8 to 15 cm tall, green grass cover of uniform height, actively growing, completely shading the ground and not short of water" (Doorenbos and Pruitt, 1977: 1). The *potential evapotranspiration*, which is the evapotranspiration that "would occur from a large area completely and uniformly covered with growing vegetation which has access to an unlimited supply of soil water" (Dingman, 2002: 232), of a crop growing under the same conditions as the reference crop is calculated by multiplying the reference crop evapotranspiration (E_r) by a crop coefficient k_c . The value of k_c depends on the stage of growth of the crop and range from 0.2 to 1.3 (Doorenbos and Pruitt, 1977: 35). The *actual evapotranspiration* (E_t) is calculated by multiplying the potential evapotranspiration by a *soil coefficient* k_s which takes into account the condition of the evaporative soil surface. Values of k_s range from 0 to 1. Thus the actual evapotranspiration E_t is calculated by:

$$E_t = k_s k_c E_r \quad (6-25)$$

where k_s is the soil coefficient, k_c is the crop coefficient and E_r is the reference crop evapotranspiration.

In practice, the method used to calculate the reference crop evapotranspiration depends on the data available at the site. Following Jenson et al. (1990: 80), the methods used to calculate E_r can be classified according to the data required. These are categorized as *temperature-based methods*, which uses air temperature and day length, *radiation-based methods*, which use net radiation and air temperature, *combination methods*, which use net radiation, air temperature, wind speed and relative humidity, and the *pan method*, which is based on the evaporation from an open water pan with modifications depending on wind speed, temperature and humidity.

The combination method is based on the Penman-Monteith Equation (Monteith, 1965) and is a modification of the original equation developed by Penman (1948). It is considered as the most “complex and physically realistic” (Liu and Todini, 2002) method for the calculation of actual evapotranspiration. However, the data required to support such a model is not extensively available and almost never exists in real-time. According to Liu and Todini (2002), the need for an extremely accurate expression for the calculation of evapotranspiration losses is not necessary in the rainfall-runoff process provided the integral effect is preserved. Thus in the TOPKAPI model, as applied by Liu and Todini (2002), evapotranspiration is calculated based on a simplified approach of the *radiation method* of Doorenbos and Pruitt (1977). It is the intention to test this concept with the application of the TOPKAPI model in South Africa in a follow-up study.

(b) Radiation method

The radiation method is suggested for areas where the measured climatic data include air temperature, and sunshine hours, cloud cover or radiation levels (Doorenbos and Pruitt, 1977: 8). A general knowledge of the levels of humidity and wind are also required. The relationship for the reference crop evapotranspiration E_r (in $\text{mm}\cdot\text{day}^{-1}$), based on the radiation method for each grid cell, is given as:

$$E_r = C_v W_{ta} R_s \quad (6-26)$$

where C_v is an adjustment factor which is obtainable from tables (Doorenbos and Pruitt, 1977: 14) as a function of the mean relative humidity and the mean daytime (07:00-19:00) wind speed at a 2 m height above the soil surface; W_{ta} is a compensation factor which is dependent on temperature and altitude for which tabulated values also exist (Doorenbos and Pruitt, 1977: 13); and R_s (in $\text{mm}\cdot\text{day}^{-1}$) is the measured short

wave solar radiation that reaches the earth's surface. R_s can be measured directly, but these data are usually not easily available for the area of investigation. R_s is dependent on the radiation received at the top of the atmosphere (R_a in mm.day⁻¹), and the transmission of this radiation through the atmosphere, which is dependent on cloud cover. Thus $R_s = (0.25 + 0.5n/N)R_a$, where n/N is the ratio of actual hours of sunshine to the maximum possible hours of sunshine in a day. R_a is dependent on latitude and the time of the year, for which tabulated values are available in Doorenbos and Pruitt (1977: 12), and tabulated mean monthly values of N , as a function of latitude, are also available in Doorenbos and Pruitt (1977: 13).

In the above expression for R_s , values of actual hours of sunshine n are not usually readily available over the individual cells of a catchment. Thus Todini (1996) sought an empirical relationship between the reference crop evapotranspiration E_r and W_{ta} , the compensation factor, the mean recorded temperature of the month T_m and the maximum number of sunshine hours N . The result is:

$$E_r = \alpha + \beta \cdot N \cdot W_{ta} \cdot T_m \quad (6-27)$$

where α and β are regression coefficients which are to be estimated for each grid cell. Eq. 6-27 is structurally similar to Eq. 6-26, the radiation-method equation, except in this instance air temperature is taken as an index of radiation and a constant has been added. According to Todini (1996), the relationship developed in Eq. 6-27 is linear in temperature and permits the disaggregation of the monthly results on a daily or sub-daily (hourly) basis. Thus the reference crop evapotranspiration E_r is expressed in mm. Δt^{-1} , and T_m , the area's mean air temperature, is averaged over Δt .

The compensation factor W_{ta} , which is dependent on the long term mean monthly air temperature and the altitude of each grid cell, is tabulated in Doorenbos and Pruitt (1977: 13). Alternatively, W_{ta} can be approximated by a fitted parabola as shown in Todini (1996).

For reasons of limited data availability, a different method for the spatial estimation of evapotranspiration is proposed here. This would be the use of calculated estimates of potential evapotranspiration, which can be made at Automatic Weather Stations (AWS) throughout South Africa (using for example the Penman-Monteith Equation), since the meteorological parameters required for such a computation are measured there.

Thereafter, through the use of the previous day's forecast of these variables via a Numerical Weather Prediction (NWP) model, these values can be spatially interpolated for the areas of interest. The establishment of this proposed method is planned for a follow up study which is currently being undertaken under a Water Research Commission Project (K5/1683; Soil moisture from satellites: Daily maps over RSA, for flash flood forecasting, drought monitoring, catchment management and agriculture). The use of actual evaporation measurements from Evaporation Pans located throughout South Africa can also be used in this instance. However, as attractive as such data might be in the application of the model, the data are notoriously inaccurate and biased and should not be used (Everson, 1999).

6.1.5. Moisture accounting in each cell

Moisture accounting in each cell is regulated by the soil store insofar as it experiences precipitation input and evapotranspiration losses directly and accepts inflows from the soil and overland stores of upstream cells. Furthermore, through horizontal subsurface flow, the soil store directly feeds the channel store of that cell (a proportionate amount) and, upon saturation, activates and feeds the overland store. The overland store also feeds the channel store of that cell a proportionate amount of flow and, together with the soil outflow, discharges to the soil store of a downstream cell. The outflow from the channel store feeds the downstream channel store directly.

In the TOPKAPI model, the grid cells are connected together in a tree shaped network. The flow of water through this network is characterised by a single preferential downstream direction in each cell, starting from "source cells" (cells without upstream contributors) downward toward the catchment outlet. The preferential direction is evaluated according to a neighbourhood relationship from a DEM of the catchment and is based on the principle of minimum energy cost (Band, 1986). This method takes into account the maximum elevation difference between the active cell and the four surrounding cells connected along the edges of the active cell. The flow path from an active cell to an edge cell is assigned in the direction of maximum slope, in either a north, south, east or west direction. Thus flow paths to the four cells diagonally adjacent to the corners of an active cell are ignored in this method. As depicted in Fig. 6-7, an active cell may have up to three contributing cells but may only feed a single downstream cell in one of the four cardinal directions.

The intra-cell operations, together with the inter-cell flows, can best be explained by examining a generic catchment consisting of three typical cells. This scenario is depicted in Fig. 6-8 where *Cell 1* (source cell) flows into *Cell 2* which in turn flows into *Cell 3*. Cell 1 is classified as a *hillslope cell*, where all surface flow is of the overland type, while the two latter cells consist of channel flow as well as overland flow. This distinction is important since not all cells have channelled flow. The classification is based on the minimum threshold area required to initiate a channel and is made from an analysis of the digital elevation model (DEM) of the catchment.

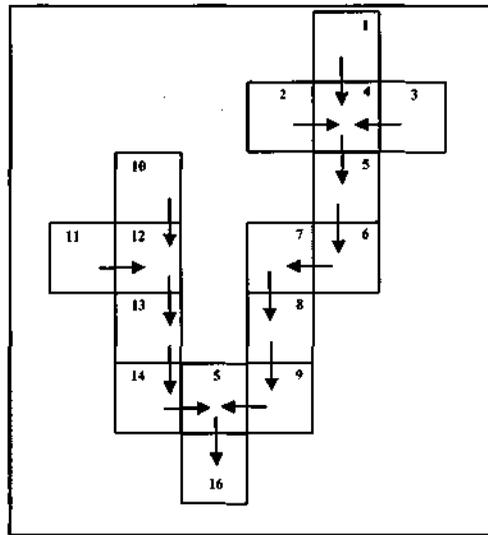


Figure 6-7. The tree-shaped network form by the cells of the TOKAPI model (from MUSIC Final Report, 2004: 53).

Within each cell, the storages operate in the following manner (with reference to Fig. 6-8). The soil store of Cell 1 receives input from incident precipitation uniformly distributed over the time step Δt . Evapotranspiration losses occur from the soil store and are subtracted as a lumped amount from the intermediate soil moisture storage $V'_{s_i}(t_o + \Delta t)$ at the end of each time step ($t_o + \Delta t$). The intermediate soil moisture volume $V'_{s_i}(t_o + \Delta t)$ is a transitional calculation step and is the solution of the non-linear differential reservoir equation (Eq. 6-17) at time $t_o + \Delta t$. Thus the actual soil moisture volume $V_{s_i}(t_o + \Delta t)$ stored in Cell 1 at the end of the time interval ($t_o + \Delta t$) results from the subtraction of evapotranspiration losses (E_a), incurred over the interval, from the intermediate soil moisture volume. This computation is shown in Eq. 6-28 for two cases, i.e. before soil saturation and after soil saturation, where in the latter case the intermediate soil moisture storage is the saturated soil moisture volume V_{sm_1} . This actual soil moisture storage is computed as:

$$V_{s_i}(t_o + \Delta t) = \begin{cases} \text{Before saturation : } V'_{s_i}(t_o + \Delta t) - E_a X^2 \\ \text{After saturation : } V_{sm_i} - E_a X^2 \end{cases} \quad (6-28)$$

where $V_{s_i}(t_o + \Delta t)$ is the actual soil moisture volume (in m^3) stored in Cell 1 at time $t_o + \Delta t$, $V'_{s_i}(t_o + \Delta t)$ is the intermediate soil store and results from the discrete solution of the non-linear differential reservoir equation (Eq. 6-14) at time $t_o + \Delta t$, $E_a X^2$ is the volumetric evapotranspiration losses over the time interval Δt and cell area X^2 and V_{sm_i} is the saturated soil moisture storage of Cell 1 which is computed from the soil properties of saturated soil moisture content θ_s and residual soil moisture content θ_r , i.e. $V_{sm_i} = V_s \cdot (\theta_s - \theta_r) = (X \cdot X \cdot L_i) \cdot (\theta_s - \theta_r)$.

The regulating function of the soil store is shown in Eq. 6-29 and is explained as follows. Once the soil store becomes saturated, all precipitation received by it becomes precipitation excess (or saturation excess) e_o which becomes the input to and hence activates, overland storage. The algorithm for this computation, for the soil store of Cell 1 in Fig. 6-8, is given by the following equation:

$$e_o = \begin{cases} \text{Before saturation : } V'_{s_i}(t_o + \Delta t) - V'_{s_i}(t_o + \Delta t) = 0 \\ \text{After saturation : } V'_{s_i}(t_o + \Delta t) - V_{sm_i} \end{cases} \quad (6-29)$$

where e_o is the saturation excess (in m^3) exfiltrated over the time interval Δt , $V'_{s_i}(t_o + \Delta t)$ is the intermediate soil store and results from the discrete solution of the non-linear differential reservoir equation (Eq. 6-14) at time $t_o + \Delta t$ and V_{sm_i} is the saturated soil moisture storage of Cell 1.

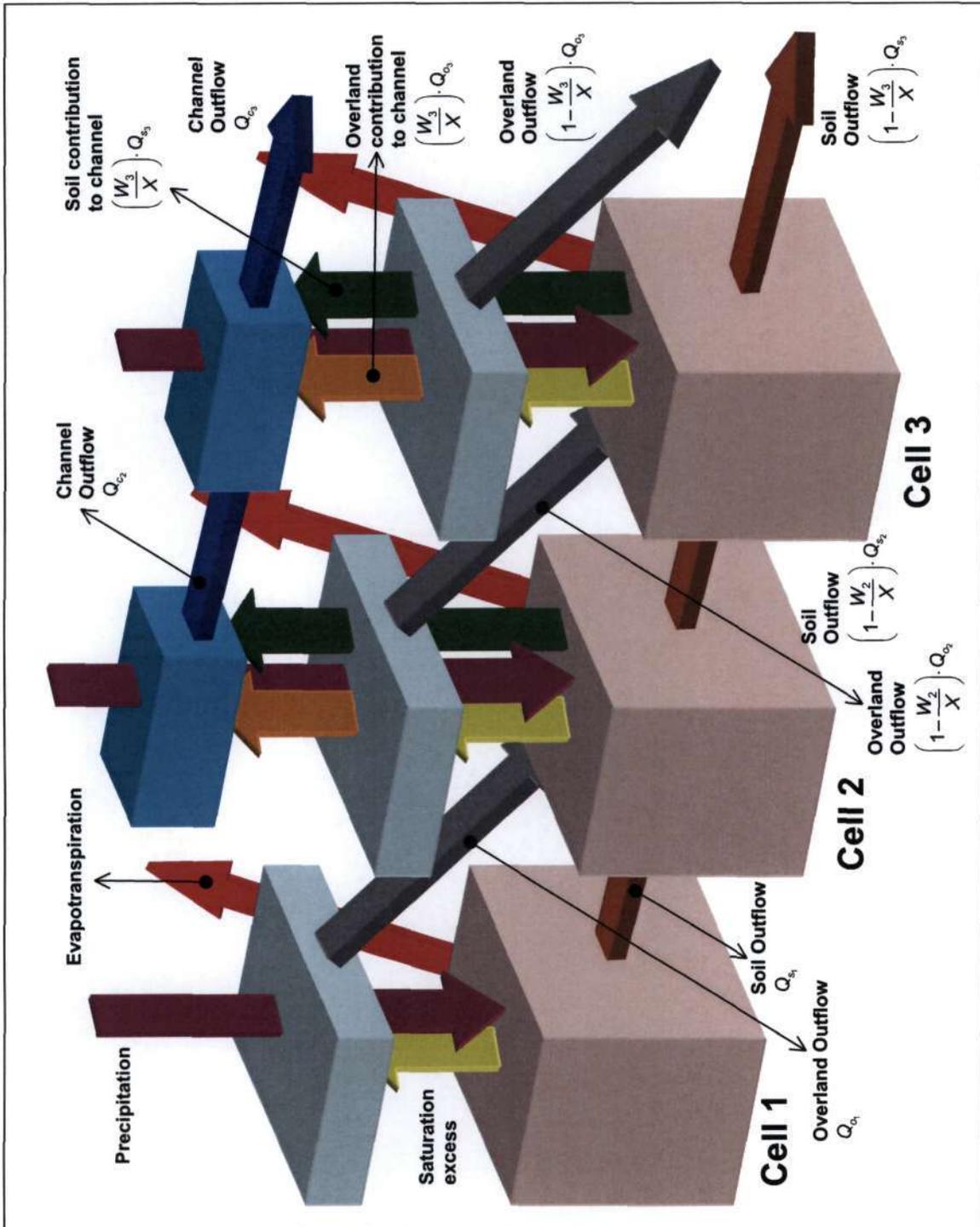


Figure 6-8. A generic catchment consisting of three cells showing the intra- and inter-cell operations of the TOPKAPI model.

The outflow from the soil store of Cell 1 (S1) will begin to flow into the soil store of Cell 2 (S2) once S1 has any moisture stored within it. This outflow is computed from the following equation:

$$Q_{s_1} = \frac{\frac{C_{s_1} X}{X^{2\alpha_s}} V_{s_1}^{\alpha_s}(t_o) + \frac{C_{s_1} X}{X^{2\alpha_s}} V_{s_1}^{\alpha_s}(t_o + \Delta t)}{2} \quad (6-30)$$

where Q_{s_1} is the average soil outflow (in m^3s^{-1}) of Cell 1 over the time interval Δt .

$\left(\frac{C_{s_1} X}{X^{2\alpha_s}} V_{s_1}^{\alpha_s}(t_o)\right)$ is the outflow from Cell 1 at the beginning of the time interval t_o and

$\left(\frac{C_{s_1} X}{X^{2\alpha_s}} V_{s_1}^{\alpha_s}(t_o + \Delta t)\right)$ is the outflow at the end of the time interval $(t_o + \Delta t)$. Hence the

average outflow over the time interval Δt from Cell 1, Q_{s_1} , becomes the uniformly distributed input for the soil store of Cell 2 (S2) in that time interval. Q_{s_1} will reach a maximum at the saturation of S1, and at this point in time, no more infiltration takes place and all incident precipitation becomes precipitation excess (or saturation excess e_o). The saturation excess computed from Eq. 6-28 is calculated as an average at the end of the time interval t_o , i.e. at time $t_o + \Delta t$, and becomes the average input for the overland storage over the time interval Δt . Outflow from the overland store of Cell 1 (O1) will flow into the soil store of the next cell S2 and will infiltrate into S2 directly, unless S2 is saturated from the previous time interval. The overland outflow from O1 to S2 is computed in a similar manner to the soil outflow from Cell 1 (Eq. 6-30):

$$Q_{o_1} = \frac{\frac{C_{o_1} X}{X^{2\alpha_o}} V_{o_1}^{\alpha_o}(t_o) + \frac{C_{o_1} X}{X^{2\alpha_o}} V_{o_1}^{\alpha_o}(t_o + \Delta t)}{2} \quad (6-31)$$

where Q_{o_1} is the average overland outflow (in m^3s^{-1}) of Cell 1 over the time interval Δt .

Continuing with the flow processes of Cell 2, the soil and overland components of this cell operate as for Cell 1, except that Cell 2 has a channel component. As such, the soil and overland components of Cell 2 (S2 and O2 respectively) will contribute some of their outflow to the channel of Cell 2 (C2) as well as to the soil store of the downstream cell. The amount that S2 and O2 contribute to C2 is proportional to the ratio of the width of the channel of Cell 2 (W_2) to the overall width of the cell (X -dimension of the cell). This is understood from the fact that $r_{c_2} \cdot X \cdot W$ from Eq. 6-23, which is the input term of the non-linear differential reservoir equation describing channel storage, is derived

from multiplying $r_{c2} \cdot X \cdot X$ (in m^3/s) by the ratio of W_2/X to obtain $r_{c2} \cdot X \cdot W_2$ for Cell 2. The term $r_{c2} \cdot X \cdot X$ comprises the sum of the overland Q_{o2} and soil Q_{s2} outflows for Cell 2, i.e. $r_{c2} \cdot X \cdot X = Q_{s2} + Q_{o2}$. Thus, a proportion equal to W_2/X of the soil and overland outflow of Cell 2 feeds the channel of Cell 2, and therefore a proportion equal to $(1-W_2/X)$ of the soil and overland flow of Cell 2 contribute to Cell 3 downstream. [This partitioning was not obvious from the original publications of the TOPKAPI model and had to be interpolated here].

The outflow from the channel of Cell 2 (C2) flows into the downstream channel of Cell 3 (C3) and is computed in a similar manner to the outflow from the respective soil and overland stores of the cell, i.e. it is computed as an average over each time step Δt :

$$Q_{c2} = \frac{\frac{C_{c2} X}{X^{2\alpha_c}} V_{c2}^{\alpha_c}(t_o) + \frac{C_{c2} X}{X^{2\alpha_c}} V_{c2}^{\alpha_c}(t_o + \Delta t)}{2} \quad (6-32)$$

where Q_{c2} is the average channel outflow (in m^3s^{-1}) leaving Cell 2 over the time interval Δt .

In a similar manner, the moisture balance for the soil, overland and channel components for any generic cell in a catchment operate in the method explained. The subsequent outflows from the representative stores to the downstream cells also operate in the manner explained. All cells of a catchment operate in this manner and the outflows from each cell are drained downwards toward the catchment outlet to give the overall outflow from the catchment.

6.1.6. Solution of the non-linear differential equations

The solution of the non-linear differential reservoir equations, describing the rate of change of moisture storages in the soil store (Eq. 6-14), overland store (Eq. 6-21) and channel store (Eq. 6-23), require complicated computational procedures. This aspect of physically-based distributed catchment modelling has in the past rendered the feasibility of such models for real-time flood-forecasting impracticable. However, with current computer power, this difficulty is no longer a limitation to this type of modelling.

There are two categories of methods that offer solutions to these equations, exact methods (analytical solutions) or approximate methods (numerical solutions). Ideally,

an analytical solution for the governing equation would be preferred, as this would reduce the computational cost of such models (which would be an added advantage for flood forecasting), however in practice most non-linear equations cannot be solved exactly. Eqs. 6-14, 6-21 and 6-23 can be written in a generalised non-linear differential form (after Liu and Todini, 2002):

$$\frac{dy}{dt} = a - by^c \quad (6-33)$$

where y represents the volume term, a the input term, b the multiplier and c the exponent of the generic equations (a , b , and c are all constant during Δt). If the value of the exponent c is 1, then Eq. 6-33 reduces to a linear form for which an analytical solution exists. However, in the TOPKAPI application of these equations, values of c range from $5/3$, which is the exponent derived from Manning's Formula for overland and channel flow, and between 2 and 4 (as suggested in Liu and Todini, 2002) for the soil (subsurface) flow. A numerical solution for the non-linear differential equations can be achieved through a variable step fifth-order Runge-Kutta algorithm due to Cash and Carp (1990) (MUSIC Final Report, 2004: 56). However, Liu and Todini (2002) present a quasi-analytical solution for Eq. 6-33 and suggest that this can reduce the computation time by one order of magnitude (MUSIC Final Report, 2004; 56). The following subsection explains the quasi-analytical solution offered by Liu and Todini (2004) which is followed by a discussion of their offered solution.

(a) Quasi-analytical solution

The quasi-analytical solutions offered by Liu and Todini (2002) are presented for three cases as there are different solutions for each case. Following from the general form of the non-linear equation given in Eq. 6-33, the three cases are based on the values of the exponent c and the constant a :

- (1) for $1 \leq c \leq 2$; for the solution of the overland and channel reservoir differential equation, where $c=5/3$,
- (2) for $c > 2$; for the solution of the soil reservoir differential equation, and
- (3) for $a=0$; when the input term in the time step is zero, independent of c .

Case 1: In Eq. 6-33, the term y^c can be approximated by a second order polynomial, i.e. $y^c = \beta y^2 + \alpha y = y(\alpha + \beta y)$, where the parameters α and β are fitted by a least squares method. Thus Eq. 6-33 can be approximately written as:

$$\begin{aligned} \frac{dy}{dt} &= a - by(\alpha + \beta y) \\ &= a - b\alpha y - b\beta y^2 \\ &= (-b\beta) \left(-\frac{a}{b\beta} + \frac{\alpha}{\beta} y + y^2 \right) \end{aligned} \quad (6-34)$$

Putting $A = -b\beta$, $B = \alpha/\beta$ and $C = -a/b\beta$, after rearranging, Eq. 6-34 reduces to:

$$\begin{aligned} \frac{dy}{dt} &= (A)(C + By + y^2) \\ \Rightarrow \frac{dy}{y^2 + By + C} &= (A)dt \end{aligned} \quad (6-35)$$

If p_1 and p_2 are the two roots of the equation $y^2 + By + C = 0$, then $(y - p_1)(y - p_2) = 0$ and:

$$p_1 = \frac{-B + \sqrt{B^2 - 4C}}{2} \quad (6-36)$$

$$p_2 = \frac{-B - \sqrt{B^2 - 4C}}{2} \quad (6-37)$$

where $p_1 \geq 0$ and $p_2 \leq 0$. The left hand side of Eq. 6-35 can be written and solved for $y_{(t+x)}$ as:

$$\int_{y_i}^{y_{i+x}} \frac{dy}{y^2 + By + C} = \int_{y_i}^{y_{i+x}} \left(\frac{1}{p_1 - p_2} \right) \left(\frac{1}{y - p_1} - \frac{1}{y - p_2} \right) dy = \int_{t_0}^t (A) dt \quad (6-38)$$

Then, the exact solution to Eq. 6-38 [the derivation was not given in Liu and Todini (2002) and was resolved in this study] is derived, step-wise, as follows:

$$\begin{aligned} \frac{1}{\rho_1 - \rho_2} \left\{ \int_{y_t}^{y_{t+\Delta t}} \left(\frac{1}{y - \rho_1} \right) dy - \int_{y_t}^{y_{t+\Delta t}} \left(\frac{1}{y - \rho_2} \right) dy \right\} &= A(t - t_0); \\ [\ln(y - \rho_1) - \ln(y - \rho_2)]_{y_t}^{y_{t+\Delta t}} &= A(\Delta t)(\rho_1 - \rho_2); \\ \ln \left(\frac{y_{t+\Delta t} - \rho_1}{y_{t+\Delta t} - \rho_2} \right) - \ln \left(\frac{y_t - \rho_1}{y_t - \rho_2} \right) &= A(\Delta t)(\rho_1 - \rho_2); \\ \ln \left(\frac{y_{t+\Delta t} - \rho_1}{y_{t+\Delta t} - \rho_2} \times \frac{y_t - \rho_2}{y_t - \rho_1} \right) &= A(\Delta t)(\rho_1 - \rho_2); \\ \frac{y_{t+\Delta t} - \rho_1}{y_{t+\Delta t} - \rho_2} = \frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \}; \\ y_{t+\Delta t} - \rho_1 &= (y_{t+\Delta t} - \rho_2) \left(\frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \} \right); \end{aligned}$$

$$\text{Finally: } y_{t+\Delta t} - (y_{t+\Delta t} - \rho_2) \left(\frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \} \right) = \rho_1$$

Now, let $\left(\frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \} \right) = \gamma$, a constant from the point of view of the

integration, then $y_{t+\Delta t} - y_{t+\Delta t}\gamma + \rho_2\gamma = \rho_1$ and the solution to this is given in Eq. 6-39.

[The analytical solution to Eq. 6-33, presented here as Eq. 6-39, is dissimilar to the final equation (their Eq. A.10) offered by Liu and Todini (2002). The final steps in the derivation of Eq. A.10 were not shown and could not be comprehended from the literature. The use of Eq. A.10 in testing the model (see Chapter 7) did not work and hence Eq. 6-39 is the analytical solution that is used in this study.]

$$y_{t+\Delta t}(1 - \gamma) = \rho_1 - \rho_2\gamma;$$

$$y_{t+\Delta t} = \frac{\rho_1 - \rho_2\gamma}{1 - \gamma};$$

$$\text{So: } y_{t+\Delta t} = \frac{\rho_1 - \rho_2 \left(\frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \} \right)}{1 - \left(\frac{y_t - \rho_1}{y_t - \rho_2} \times \exp \{ A(\Delta t)(\rho_1 - \rho_2) \} \right)} \quad (6-39)$$

Case 2: For horizontal subsurface flow in the soil store, the value of the exponent c ranges from 2 to 4 (Liu and Todini, 2002). By making the substitution $u = y^{-(c-1)}$, then:

$$\begin{aligned} \frac{du}{dy} &= \frac{-(c-1)}{y^c} \\ \Rightarrow \frac{dy}{y^c} &= \frac{du}{-(c-1)} \end{aligned} \quad (6-40)$$

By substituting this result into Eq. 6-33, the following is obtained:

$$\begin{aligned} \frac{dy}{dt} \frac{1}{y^c} &= \frac{a}{y^c} - b \\ \Rightarrow \frac{du}{dt} \frac{1}{-(c-1)} &= \frac{a}{y^c} - b \\ \Rightarrow \frac{du}{dt} &= b(c-1) - \frac{a(c-1)}{y^c} \end{aligned} \quad (6-41)$$

It can be shown that $\frac{1}{y^c} = u^{\frac{c}{c-1}}$ from the initial substitution of $u = y^{-(c-1)}$, and thus Eq.

6-41 reduces to:

$$\frac{du}{dt} = b(c-1) - a(c-1)u^{\frac{c}{c-1}} \quad (6-42)$$

Since the term $\frac{c}{c-1}$ falls into the range of 1 to 2 where c is in the range of 2 to 4, the

term $u^{\frac{c}{c-1}}$ can be approximated by the second order polynomial in the method described for Case 1, i.e. $u^{\frac{c}{c-1}} = u(\alpha + \beta u)$. In this case:

$$\frac{du}{dt} = A(u^2 + Bu + C) \quad (6-43)$$

where now, $A = -a(c-1)\beta$, $B = \frac{\alpha}{\beta}$ and $C = -\frac{b}{a\beta}$. The parameter a represents the input terms, b the multiplier, c the exponent, and α and β the fitted variables. The solution of Eq. 6-43 can be accomplished in the same manner as for Case 1.

Case 3: In this case the term a is zero. This situation arises when there is no more inflow into a cell, either from precipitation input or from upstream cells. Therefore, Eq. 6-33 reduces to the following:

$$\begin{aligned} \frac{dy}{dt} &= by^c \\ \Rightarrow \frac{dy}{y^c} &= (-b) dt \end{aligned} \tag{6-44}$$

This is easily integrated and the difference equation takes on the following non-linear form:

$$y_{t+\Delta t} = \left[y_t^{(1-c)} + b(c-1)(\Delta t) \right]^{\frac{1}{(1-c)}} \tag{6-45}$$

where $y_{t+\Delta t}$ is the y at time $t+\Delta t$ and y_t is the initial value of y at each time step.

(b) A comment on the analytical solution

The key aspect to the derivation of the analytical solution offered for the non-linear differential equation is in the approximation of the non-linear term, y^c , by a second order polynomial. For the overland and channel stores, this approximation is

$y^c = y(\alpha + \beta y)$ and for the soil store the approximation is $u^{\frac{c}{c-1}} = u(\alpha + \beta u)$, where α and β are fitted variables and $u = y^{-(c-1)}$ is an integration substitution. Since the exponent c for the overland and channel store equals $\frac{5}{3}$, then the best fit of the variables α and β are if $\alpha = 0$ and $\beta = \frac{1}{y^{\frac{1}{3}}}$. Thus $\frac{y^{\frac{5}{3}}}{y} = \alpha + \beta y \Rightarrow y^{\frac{2}{3}} = 0 + \frac{1}{y^{\frac{1}{3}}} y$ and

$y^{\frac{2}{3}} = y^{\frac{2}{3}}$. For the soil store, if $c = 2$, then $\frac{u^{2-1}}{u} = \alpha + \beta u \Rightarrow u = 0 + 1 \cdot u$ and the best fit of the variables α and β are $\alpha = 0$ and $\beta = 1$. Similarly, if $c = 3$ and 4, then the best fit

of the variables are if α equals zero in both instances and $\beta = \frac{1}{u^{1/2}}$ and $\beta = \frac{1}{u^{1/3}}$ respectively. Since $u = y^{-(c-1)}$, then for $c = 2, 3$ and 4 , $\beta = 1, y$ and y^2 respectively. This corresponds exactly to the relationship $\beta = y^{2-c}$ and $\alpha = 0$ over the full range of c from $5/3$ to 4 .

The result of this is that the variable β is now implicitly a function of y . This means that, when β is fitted at the beginning of the time step Δt , it is a function of y_t (the volume in the store at the beginning of the time step). However, as the volume changes during the time step, the value of β is still fixed to y_t and cannot match the change in volume. This setup will have the tendency to underestimate or overestimate $y_{t+\Delta t}$ (the volume at the end of the time step), depending on if the hydrograph is on the rising limb or falling limb respectively. This issue is investigated and resolved on Chapter 7 when a test application of the model reveals this “lack of fit” in the analytical solution setup.

6.1.7. Calibration

The TOPKAPI model is a physically-based model and as such all input parameters can be obtained directly from field measurements and related literature. In theory, it requires no calibration. However, Liu and Todini (2002) suggest that the calibration of parameters is still necessary because of the uncertainty of the information on topography, soil characteristics and land cover and also because of the approximations introduced by the scale of the parameters representations. They maintain that the calibration of the model is “more an adjustment” and is achieved through simple trial and error methods. The parameters are calibrated on a continuous sequence of a selected portion of historical precipitation and flow data for the catchment. The parameters are adjusted such that the observed outflow from the catchment mimics the outflow simulated using the model.

6.2. Chapter summary

This chapter was intended to introduce the TOPKAPI model for its application in the Liebenbergsvlei catchment. As such, various items of literature on the TOPKAPI model and other related subjects were assimilated, dissected and explained here as a coherent whole with the intention that the manner in which the model operates could be correctly understood. Various issues related to the model's setup were also identified, such as that of distributed evapotranspiration estimation, flow partitioning and the formulation of the analytical solution. The latter two issues are resolved in Chapter 7 where the model is implemented in a test environment. The distributed estimation of evapotranspiration is left for a follow up study. Chapter 7 also deals with the model's data requirements and the methods used to acquire and manipulate the required input data for the study catchment.

CHAPTER 7

7. THE APPLICATION OF THE TOPKAPI MODEL

Chapter 6 introduced the TOPKAPI model and explained the theory and methodology of its operations. Various issues related to the model's implementation were also identified. This chapter is intended to describe the application of the model as an operational hydrologic catchment model for real-time rainfall-runoff applications. This is achieved by firstly describing the types of data that are required by the model as well as the efforts that were put into acquiring and preparing this data to sufficiently model a catchment, namely the Liebenbergsvlei catchment. This is dealt with in Section 7.1, where the required data are sourced through geographical information systems (GIS) and remote sensing techniques, which were managed on ARCGIS™ (a GIS software platform). Section 7.2 deals with the "test" application of the model. This consisted of establishing a "four cell generic catchment" where the model was run to verify that the intricate operations of the model were correctly interpreted and implemented. The issues identified in Chapter 6 are also resolved. This was achieved through the use of a standard spreadsheet package offered by Microsoft Excel™.

The chapter ends with a summary (Section 7.3) detailing the work still required in order to completely establish the TOPKAPI model as a functioning real-time rainfall-runoff application for flood-forecasting purposes. The intended outcome of this section of the research is on laying the groundwork necessary for the establishment of the TOPKAPI model as a fully-distributed hydrologic model in the Liebenbergsvlei catchment; the latter aspect was beyond the scope of this study.

7.1. Data requirements

The TOPKAPI model requires two kinds of input data, namely static (or very slowly varying) and dynamic. The static data required are terrain data (from a digital elevation model (DEM) of the catchment), soil data and vegetation cover or landuse data. The dynamic data required are estimates of measured and/or calculated evapotranspiration and precipitation.

In terms of the static parameters, there are seven classes of input parameters that are required by the model in each grid cell, namely:

- L (the thickness of the surface soil layer, in m),
- k_s (the saturated hydraulic conductivity of this layer in ms^{-1}),
- θ_r (the residual moisture content of the soil),
- θ_s (the saturated moisture content of the soil),
- α_s (the pore-size distribution exponent for the transmissivity of the soil, which is taken as constant for all the cells in a catchment), and
- n_o and n_c (which are the surface and channel roughness coefficients respectively (in $\text{m}^{-1/3}\text{s}^{-1}$) according to Manning's Equation).

The first five classes of parameters relate to the soil and are responsible for the production of runoff. These parameters are obtainable from literature as a function of the soil type. The type of soil present in each cell is identified from a *soils map* of the catchment. The last two classes of parameters are responsible for the routing of runoff, over the hillslopes and in the channel respectively. These are also obtainable from literature as a function of the landuse or landcover properties of the cell, which is identified from a *landuse map* of the catchment.

The DEM application in the model consists of describing the topographic and geomorphologic elements of the catchment, in terms of calculating the surface slopes, areas drained, identifying the flow pathways and detecting the drainage networks. The primary source of this data and the methods of manipulation and analyses are through GIS techniques, which are described in detail in the subsections that follow. Since each pixel of a DEM forms the primary unit of the processing cells in the TOPKAPI model, it was decided to model the hydrologic processes of the Liebenbergsvlei catchment at the 1km spatial scale. It was shown in Chapter 6 that the distributed operations of the TOPKAPI model are still physically valid up to a grid cell size of 1km square (Martina, 2004: 76). Thus, it was decided to standardise on this modelling resolution as the various static input data are easily and freely available at this scale and, more importantly, distributed rainfall input (remotely sensed from radar) is only available at this accuracy in South Africa. However, it was found that a DEM resolution at this scale was too coarse to accurately identify the flow pathways and detect the drainage networks of the Liebenbergsvlei catchment and so it was necessary to resample a finer resolution DEM up to the desired scale (1km square). This, and other manipulation techniques that were necessary, are expanded on in more detail in the subsections that

follow. However, these explanations are preceded by a description of the test catchment that was the focus of this study, i.e. the Liebenbergsvlei catchment.

7.1.1. The Liebenbergsvlei catchment

The Liebenbergsvlei catchment is a sub-catchment of the Vaal drainage basin and is situated near Bethlehem in the Free State Province. The area of the catchment is approximately 4625km² and consists predominantly of dry cropland and grassland. The location of the catchment, in relation to South Africa, is shown below in Fig. 7-1 together with its quaternary sub-catchment divisions and river network (from Midgley et al., 1994), which are shown at a spatial detail of 1:250 000. The locations of the 45 telemetering raingauges and the 9 flowgauges that are found in the catchment are also shown as well as the location of MRL5 S-Band weather radar that covers the catchment. These instruments provide temporally and spatially detailed hydrologic data necessary to implement a distributed catchment model such as the TOPKAPI.

Topographic data for the Liebenbergsvlei catchment, in terms of the DEM, was sourced from HYDRO1k (1996) and DLSI (1996) for pixel resolutions of 1km and 218m square respectively. The 1km square DEM is shown below in Fig. 7-2, where this resource is derived from the geographic database of the United States Geological Survey (USGS) '30 arc-second digital elevation model of the world. This DEM is freely available from the USGS website (<http://edcdaac.usgs.gov/gtopo30/hydro>) for all the continents of the world at the 1km spatial resolution and includes other topographically derived data sets, such as stream networks and drainage basins which have been hydrologically processed for errors.

It was the aim of this research to model the Liebenbergsvlei catchment at the 1km spatial resolution and as such it was initially intended to use the HYDRO1k (1996) DEM for this purpose. However, the resolution of this DEM proved to be too coarse in terms of accurately tracing the catchment boundaries and the stream networks, as explained in Section 7.1.2. This was determined by comparing the stream network delineated from the 1km DEM with that digitised from topographic maps (from Midgley et al., 1994). The latter information was captured at a spatial detail of 1:250 000 and is shown in Fig. 7-1. Thus a finer resolution DEM was sought (DLSI, 1996) and the topographic analysis of the Liebenbergsvlei catchment that proceeded was based on this DEM. In order to maintain the chosen modelling scale of 1km square, the pixel resolution of this latter resource was "resampled" to the desired scale. This will be

made clearer in Section 7.1.2. The soils type and landuse information was obtained from GLCC (1997) and SIRI (1987) respectively. The processing of these data is covered in Section 7.1.3 and 7.1.4 respectively.

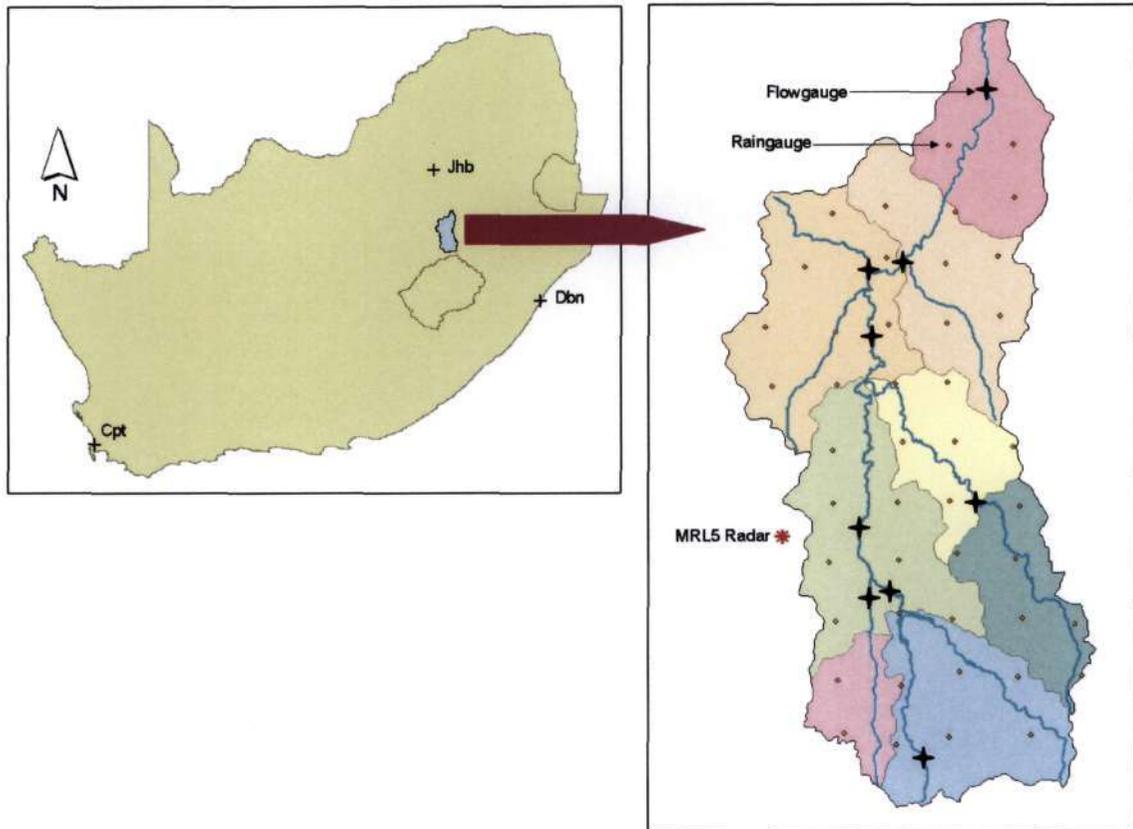


Figure 7-1. The Liebenbergsvlei catchment in South Africa showing the eight quaternary sub-catchments and the river network, which is shown at a spatial detail of 1:250 000 (Midgley et al., 1994). Also shown are the locations of the 45 telemetering raingauges, 9 flowgauges and the S-Band MRL5 weather radar.

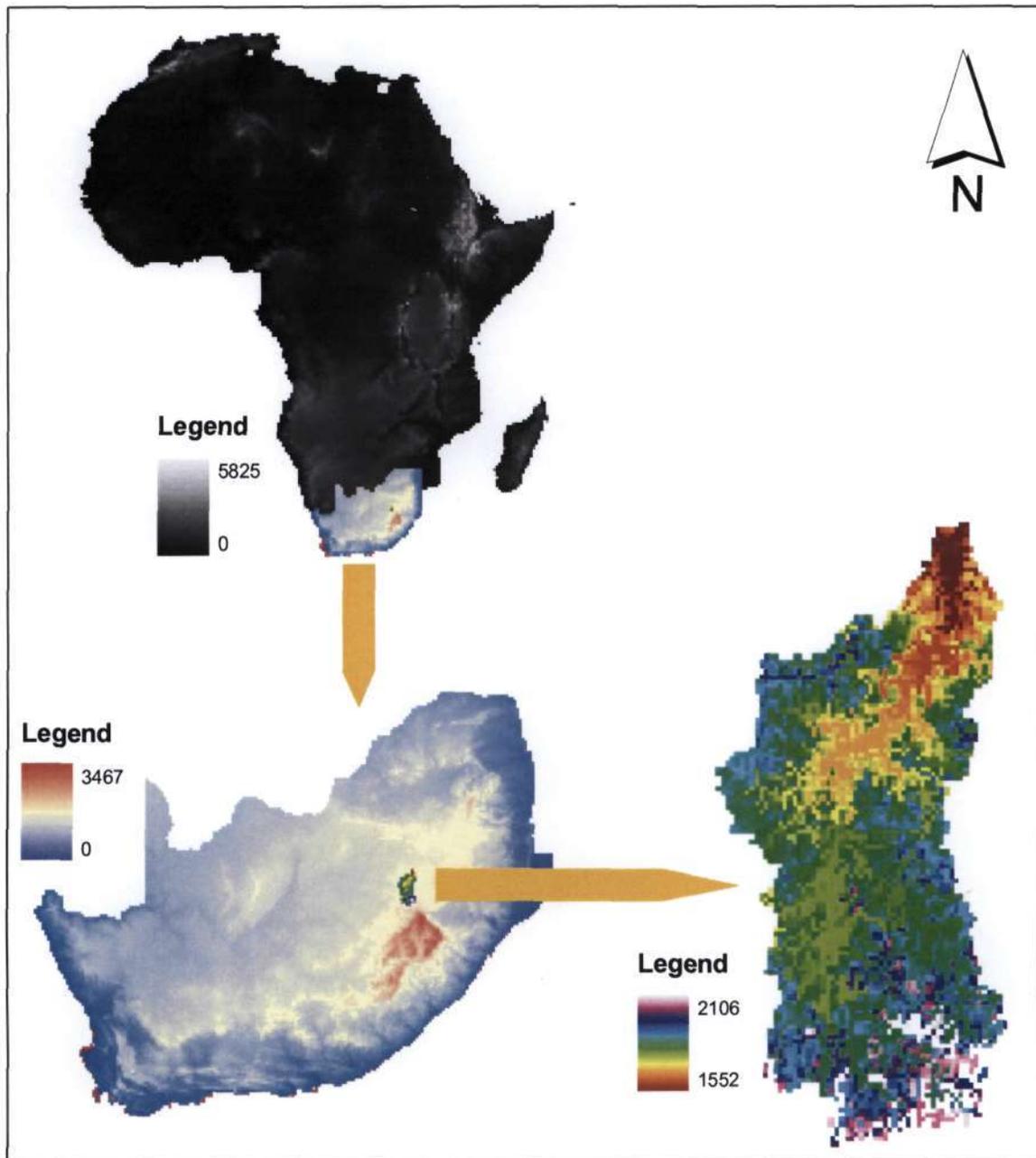


Figure 7-2. A DEM of Africa, South Africa and the Liebenbergsvlei catchment at a grid resolution of 1km square from the geographic database of the United States Geological Survey (HYDRO1k, 1996).

7.1.2. Digital Elevation Model (DEM)

(a) Background

A DEM is a raster-based description of a continuous surface and represents a grid of elevation heights, in cells or pixel format, above some datum (such as sea level). The accuracy of the DEM is highly dependent on the resolution of the grid cells, where a

coarser resolution DEM is more prone to contain errors and is not able to accurately represent surface features and permit the detection of flow pathways.

A common problem in DEM's is the occurrence of *sinks*. A sink, also referred to as a depression or pit, is a cell or area that is surrounded on all sides by higher elevation values. As such, a sink is an area of internal drainage and prevents the downslope routing of water and, unless it is an actual case such as a lake or swamp, it is an error. By contrast, a peak is a cell or area surrounded by lower elevation values and drains water away from it. Peaks are sometimes erroneous, but are more likely to be natural features. However, it is very rare that sinks are natural features and are more likely to be errors in the DEM (Mark, 1988). These errors often arise due to the sampling techniques used in processing a DEM or due to the rounding off of elevation values to integers.

In order to create an accurate representation of the flow direction, it is best to use a DEM that is free of sinks, or a *depressionless DEM*. To create such a DEM from an existing DEM, the sinks need to be filled. This is an iterative process, since the filling of a sink cell or area may create a new sink at the boundary of the filled cell, which in turn needs to be filled. A sink is filled to its outflow point, which is the minimum fill elevation required in order for water to flow out of the cell into a neighbouring cell. This can be achieved using the GIS software package, ARCGIS™.

Once a depressionless DEM is created the next step in the delineation of the stream network is the determination of the outflow direction of each cell, i.e. the direction of the steepest outflow path from an active cell to the neighbouring downstream cells. A common algorithm used for this purpose is the *D8 flow model* of O'Callaghan and Mark (1984). This method assigns the outflow from an active cell into its neighbour along one of the eight possible paths to which it could flow, i.e. the four cardinal paths and the four diagonal paths. This outflow path is defined as the path of steepest slope. Fig. 7-3 shows how this method works as applied in the ARCGIS™ environment. Given a DEM (on the top left of Fig. 7-3), a drainage *direction code* is assigned to each cell (shown on the top right of Fig. 7-3) based on the direction codes which are shown in the bottom panel. This code depends on the direction of maximum slope, which is calculated as the maximum difference in elevation divided by the horizontal distance from the centre of the active cell to the centres of the eight surrounding cells. If the maximum slopes to several cells are the same, then the neighbourhood around the active cell is enlarged until the direction of steepest slope is found. Fig. 7-4 shows the flow direction raster

computed on ARCGIS™ for the Liebenbergsvlei catchment based on the 218m square resolution DEM (DLSI, 1996); from here on this DEM will be referred to as the 200m DEM.

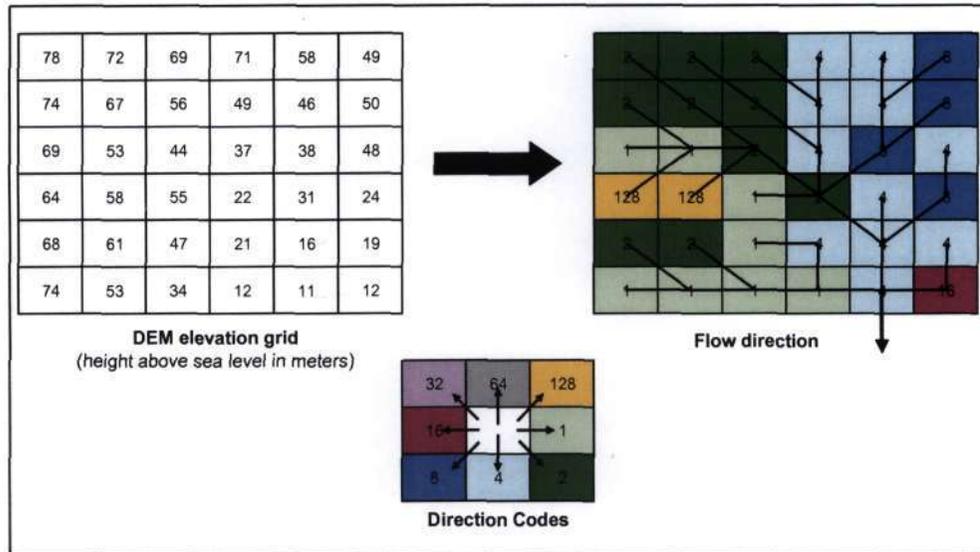


Figure 7-3. The D8 flow model of O'Callaghan and Mark (1984) as applied in ARCGIS™.

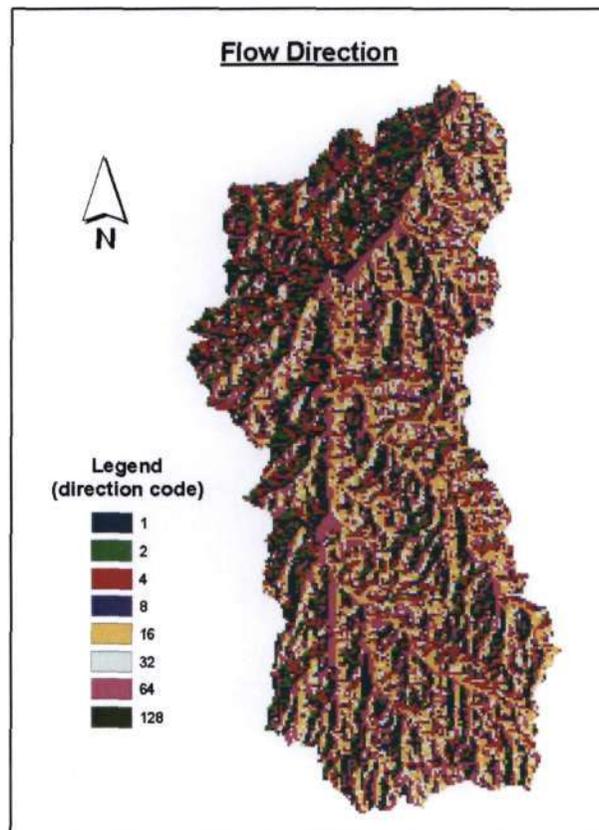


Figure 7-4. A flow direction raster showing the direction code of each cell determined using ARCGIS™ for a 200m square resolution DEM of the Liebenbergsvlei catchment.

The next step in the delineation of stream networks is to determine the number of upslope cells that contribute flow into each cell, i.e. the flow (in terms of contributing cells) accumulated in each cell. This is also achieved on ARCGIS™ with a standard tool. Fig. 7-5 shows the flow accumulation raster for the Liebenbergsvlei catchment. The colour palette indicates, for each cell, the number of upslope cells that feed it. The main trunk of the stream network is easily visible from this image.

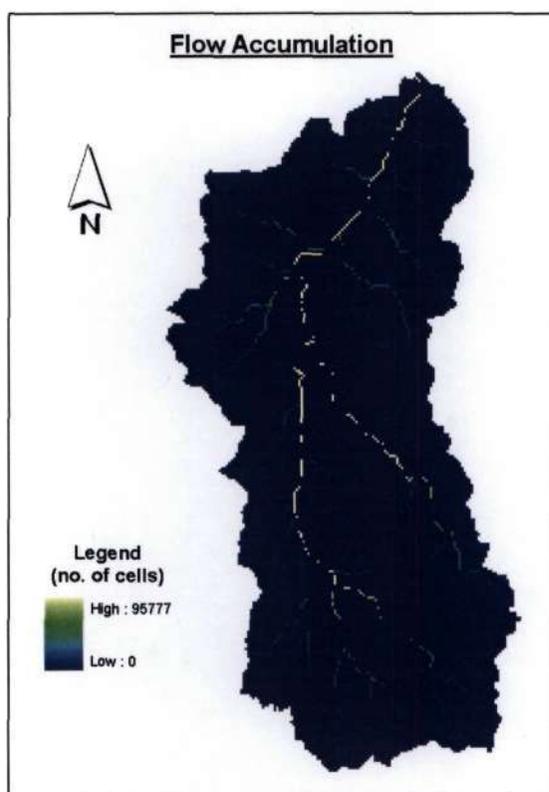


Figure 7-5. A flow accumulation raster showing the number of upslope cells flowing into each cell determined using ARCGIS™ for the Liebenbergsvlei catchment.

The final step in delineating the stream network from a DEM is to assign a threshold value to the *flow accumulation* raster, for the minimum number of upslope cells that are required to initiate a channel in an active cell. The determination of this threshold value depends on, according to Tarboton et al. (1991), climate, slope and soil characteristics. Tarboton et al. present procedures in order to “rationally select the scale at which to extract channel networks” which correspond to networks obtained through more traditional methods, such as from topographic maps or fieldwork. Fig. 7-6 shows a comparison between the stream network delineated from the 200m DEM in the manner described above (shown on the left of Fig. 7-6) and a stream network digitised from a topographic map of the catchment (shown on the right of Fig. 7-6) at a spatial scale of 1:250 000 (from Midgley et al., 1990). The threshold value chosen was 500 pixels,

which corresponds to an area of approximately 20km² for a pixel size of 200m square. The comparison shows good correspondence between the two sources of networks.

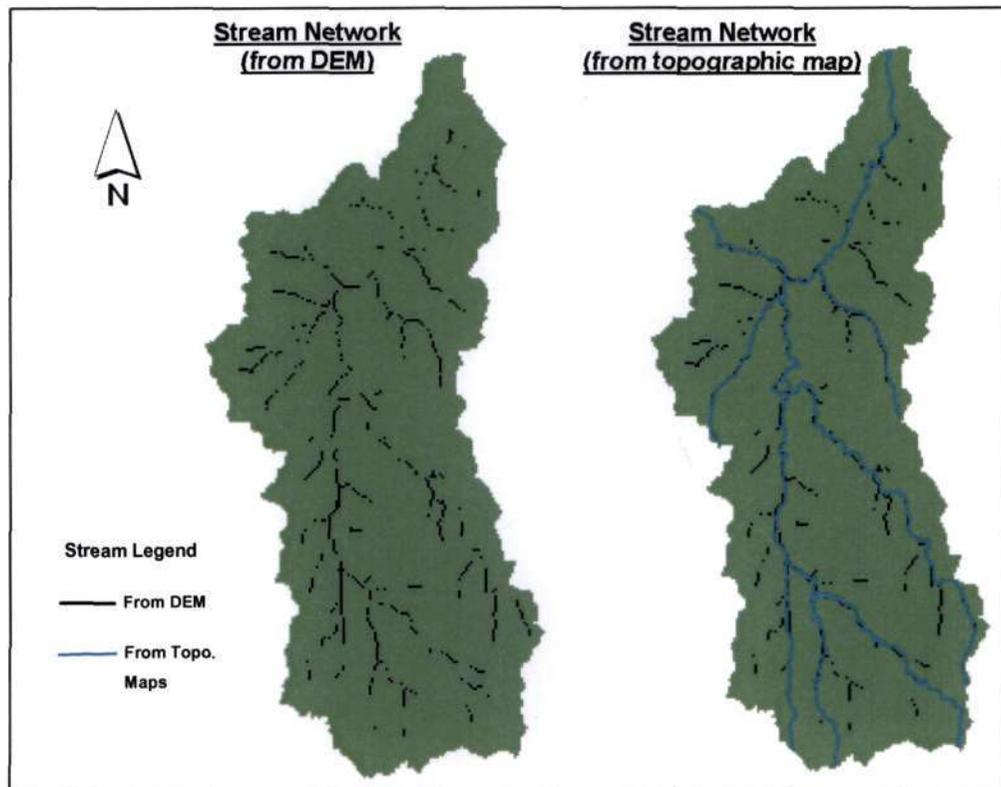


Figure 7-6. A comparison between a stream network delineated for the Liebenbergsvlei catchment from a DEM, using a threshold value of 500 cells and a stream network digitised from a topographic map (at a spatial scale of 1:250 000 from Midgley et al. (1990)). The DEM delineated network shown on the left, although appearing disjointed in the image, is continuous in reality.

Further topographic data can be computed on ARCGIS™ which are required by TOPKAPI as input. Data such as the stream orders of the delineated stream network and the surface slopes of each pixel are required. These rasters can be computed easily using inbuilt functions on ARCGIS™. The computation of these rasters will not be explained here since they are standard procedures using the GIS software. The stream order and surface slope rasters are shown and their use explained in the relevant Subsection (b) that follows. This section describes the effort that was undertaken in order to obtain TOPKAPI specific input from the DEM of the Liebenbergsvlei catchment.

(b) DEM application in the TOPKAPI Model

As mentioned in Section 7.1.1, it was required to model the Liebenbergsvlei catchment using a 1km grid cell resolution. To accurately achieve this, the 218m square pixels of the 200m DEM were transformed to 1000m (1km) square pixels. This was achieved using an inbuilt function on ARCGIS™ called “resample”. The result of this is shown in Fig. 7-7 (on the right). The 1000m resampled DEM shown in Fig. 7-7 (from here on referred to as the 1km DEM) has been processed for sinks. The “resample” function has the option of three interpolation techniques, i.e. nearest neighbour assignment, bilinear interpolation and cubic convolution. Nearest neighbour assignment uses the closest value from the cell on the input raster to assign to the new “resampled” cell on the output raster. It is appropriate for categorical data, such as landuse rasters. For continuous data, such as elevation rasters, the bilinear interpolation and cubic convolution techniques are preferred as they make use of a greater number of nearby cells (four and sixteen respectively) to compute the value of the new transformed cell. These techniques (bilinear and cubic) make use of a weighted average, based on distance from the centres of the input cells to the centre of the output cell, to compute the new “resampled” value. In this instance, the bilinear interpolation function was used to resample the 200m DEM into a 1km DEM as it was found that the cubic convolution method often gives values outside of the input range. The height range of the new resampled 1km DEM can be seen to be less than the height range of the original 200m DEM (in terms of maximum elevation). This is an output result of the resampling function, but, as will be seen further on, it does not affect the delineation of the catchment boundaries and stream networks.

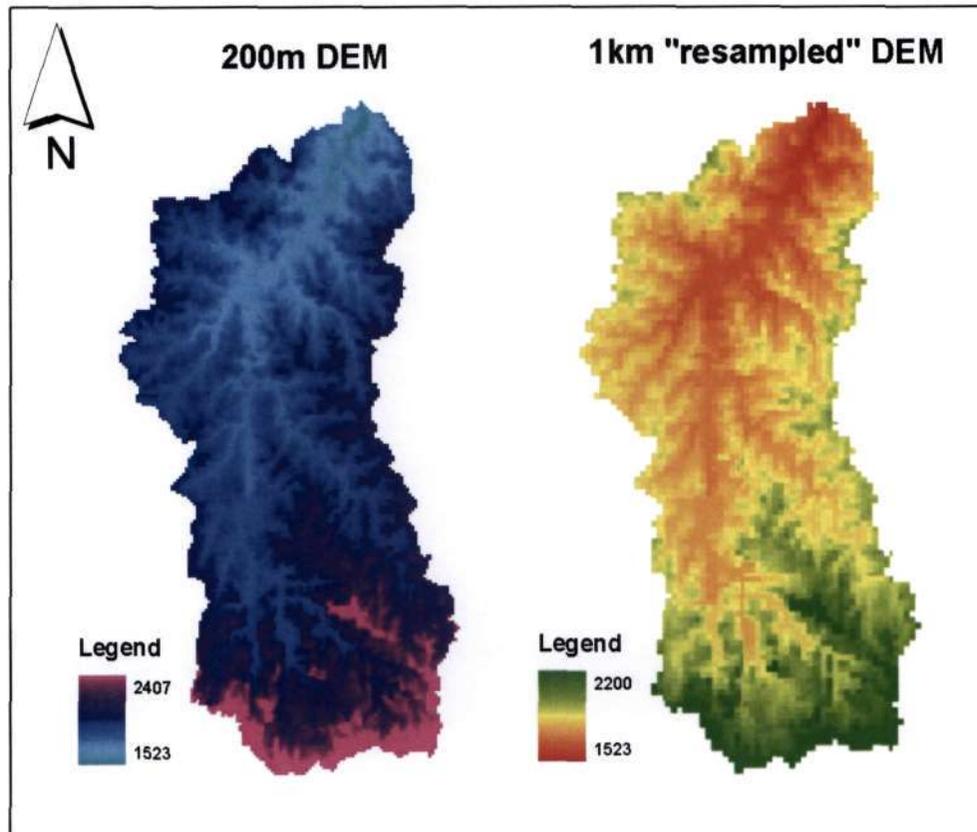


Figure 7-7. A 218m square pixel resolution DEM of the Liebenbergsvlei catchment is shown on the right (from DLSI, 1996). The pixels of this DEM were “resampled”, using an inbuilt function on ARCGIS™, to a resolution of 1000m square (shown on the left).

Once the 1km DEM was obtained, the topographical processes described in Subsection (a) above were performed in order to obtain the desired input for the TOPKAPI model. These included the surface slope raster, the determination of the outflow direction of each cell and the computation of the stream networks and stream orders. These rasters are shown and explained in the following relevant subsections. However, the determination of the outflow direction of each cell required additional manipulation techniques outside the scope of the capabilities of ARCGIS™. This was in regard to a specific requirement of the TOPKAPI model which was that the outflow path from each cell was to be limited to either a north, south, east or west direction. Thus the *D8 flow model* of ARCGIS™ was not suitable for the determination of the flow direction raster for input into the model. Therefore, additional manipulation was required with regard to this aspect of the research and is detailed below in the Subsection (i).

(i) Flow direction and flow accumulation

As defined above, a flow direction raster computes the direction of flow from each cell in the catchment. This is an important part of catchment modelling as it determines the connectivity between cells and is the first step in tracing the stream network whose authenticity is essential. On the ARCGIS™ platform, the flow directions of each cell are determined based on the D8 method, i.e. flow is assigned in any one of eight directions from an active cell, as explained earlier. The TOPKAPI model operates on the four cardinal directions only, in such a manner that an active cell may receive flow from up to three upstream cells and may only have one outflow direction to a downstream cell. Since the flow direction function on ARCGIS™ operates on the D8 method, it was necessary to create a “D4” flow direction function which would be compatible with the TOPKAPI’s requirements.

It was decided that the simplest manner in which to accomplish a D4 flow direction raster was to resolve the four diagonal directions of a D8 flow direction raster into the four cardinal directions. This is shown in Fig. 7-8, where the four diagonal direction codes (i.e. 128, 2, 8 and 32) are resolved into direction codes 1, 4, 16 and 64 respectively; those cells that had direction codes reflecting the four cardinal directions from the original D8 raster were left unchanged. The D8 flow direction raster for the Liebenbergsvlei catchment was produced using the GIS software ARCGIS™. In order to resolve a diagonal direction into either a north/south or an east/west direction, three methods were tested. The first was to resolve the diagonal direction code toward that neighbouring vertical or horizontal direction that had the lowest elevation (based on a DEM), and hence toward that neighbouring direction of steepest slope. The second option was to arbitrarily resolve the diagonal directions in a clockwise manner. Referring to Fig. 7-8 for example, direction code 128 would be resolved clockwise to direction code 1; direction code 2 would be resolved to direction code 4, and the rest of the codes would be resolved in this manner. The third option that was tested was to resolve the direction codes in an anti-clockwise manner in a similar manner to the second option. The resolving processes were accomplished through M-File Programming Functions that were written on MATLAB™. These codes are included in Appendix C.

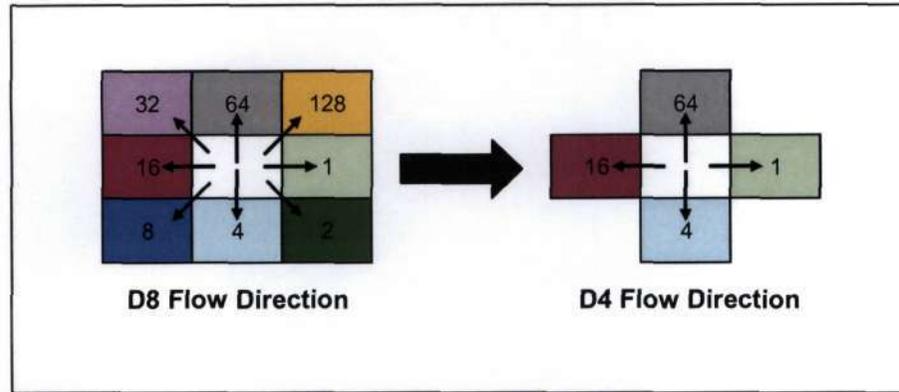


Figure 7-8. The D8 flow direction raster codes resolved into a D4 flow direction raster.

The results of all three tests had similar outcomes, in terms of the flow accumulation rasters that were produced from the flow direction rasters (determined using the three different methods explained). An example of the D4 flow direction raster of the Liebenbergsvlei catchment, output from the first test of this exercise (i.e. assigning the diagonal direction code toward that neighbouring vertical or horizontal direction that had the lowest elevation), is shown on the left of Fig. 7-9. Intuitively, the resolving of the direction codes based on lowest elevation makes more sense than resolving the codes arbitrarily in one direction. However, it is interesting to note that the different methods used displayed no apparent skill over the other. Thus, based on theoretical reasoning, it was decided to use the “steepest neighbour” method in the computation of the flow accumulation raster. The resultant flow accumulation raster, from the new D4 flow direction raster, is shown on the right of Fig. 7-9.

It is evident from Fig 7-9 that the flow accumulation raster is not continuous and hence does not accurately reflect the flow paths and stream networks. Secondly, the total number of pixels, at a resolution of 1km square, that cover the Liebenbergsvlei catchment are 4625. This number should be reflected in the accumulation raster’s legend at the catchment outlet since all cells of the catchment should contribute flow at this point. The number of cells shown on the right of Fig. 7-9 is 1947 which indicate that all cells of the catchment are not contributing flow at the outlet. These points show that the process of resolving the D8 direction codes into the four cardinal direction codes had created areas of internal drainage within the catchment. These areas of internal drainage are sinks created artificially as a result of limiting the outflow drainage directions to four. This did not occur in the D8 flow direction model of this catchment since all sinks were previously filled based on the eight surrounding cells. This situation is visually explained in Fig. 7-10, where the sinks were created as a result of

anomalous positions of direction codes (based on the D4 method) that cause flow towards each other.

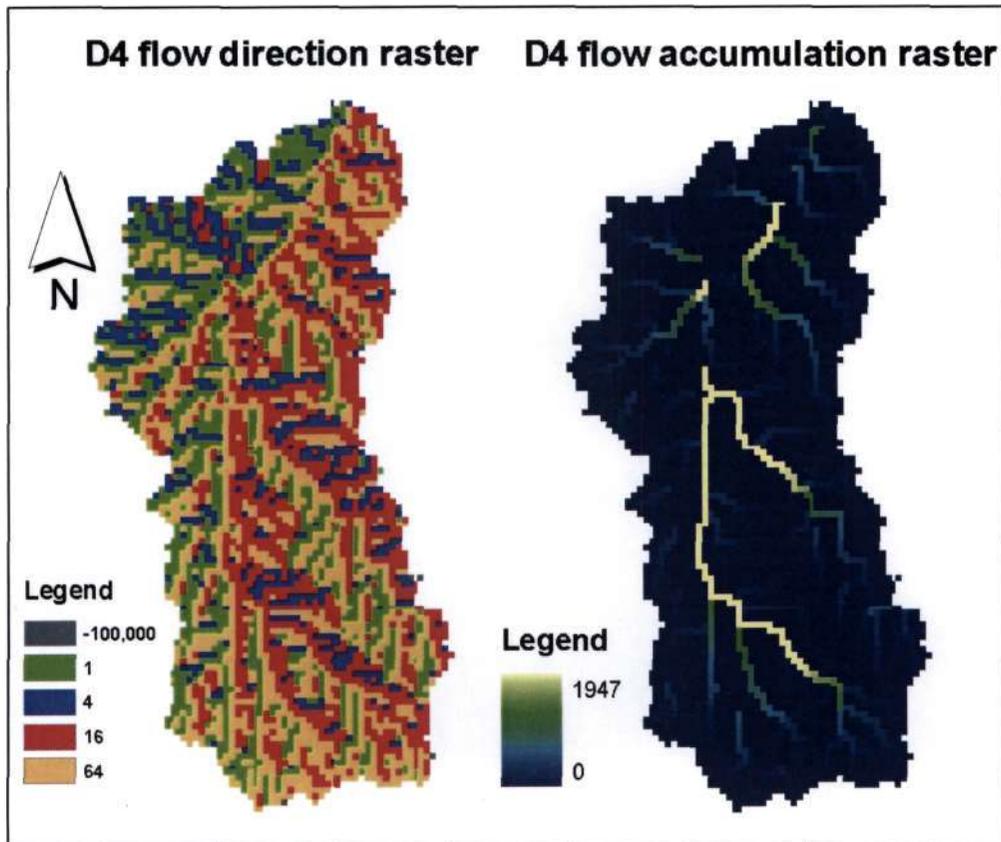


Figure 7-9. On the left is shown a D4 flow direction output as a result of resolving a D8 direction raster of the Liebenbergsvlei catchment into the four cardinal directions only. The legend of the direction raster show the four cardinal direction codes as well as an error code (-100 000) to mark those cells whose resolved directions flowed out of the catchment. On the right is shown the resulting flow accumulation raster based on the D4 method.

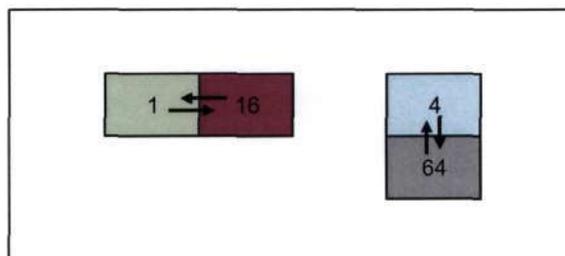


Figure 7-10. Anomalous positions for direction codes that resulted from the resolving of a D8 flow direction raster into a D4 raster. These pixels created areas of internal drainage within the catchment.

In order to solve this problem, further code was written on MATLAB™ to identify the “problem” pixels shown in Fig. 7-10 as well as those pixels along the catchment boundary that flowed outwards. This code is also included in Appendix C. A “catchment

mask”, which displayed the problem pixels, was then created on Microsoft Excel™ by importing the code’s output. The direction codes of the problem pixels were then easily rectified (by hand) based on this masking technique. The number of instances that required rectification was: 12 pixels that flowed outwards, 13 cases of pixels flowing toward each other in a north/south direction and 14 cases of pixels flowing toward each other in an east/west direction. The result of this can be seen in Fig. 7-11, where the flow accumulation raster is shown on the left, after manual checking of the direction codes was performed (in the manner explained above). The accumulation raster is overlain on the right of Fig. 7-11 with the stream network traced from a topographic map (from Midgley et al., 1994) at spatial scale of 1:250 000.

The results shown in Fig. 7-11 summarise, visually, that the processes explained above are able to trace the correct flow paths of the catchment based on a D4 method, which is a requirement of the TOPKAPI model. Secondly, the number of pixels that contribute flow to the catchment outlet is shown in the legend of Fig. 7-11 as 4625. This number is in agreement with the number of pixels contained in the Liebenbergsvlei catchment at a resolution of 1km square. It is envisaged that through more skilled programming techniques, the entire process of resolving a D8 flow direction raster to a D4 raster could easily be achieved in one script. This code has not been written in this study.

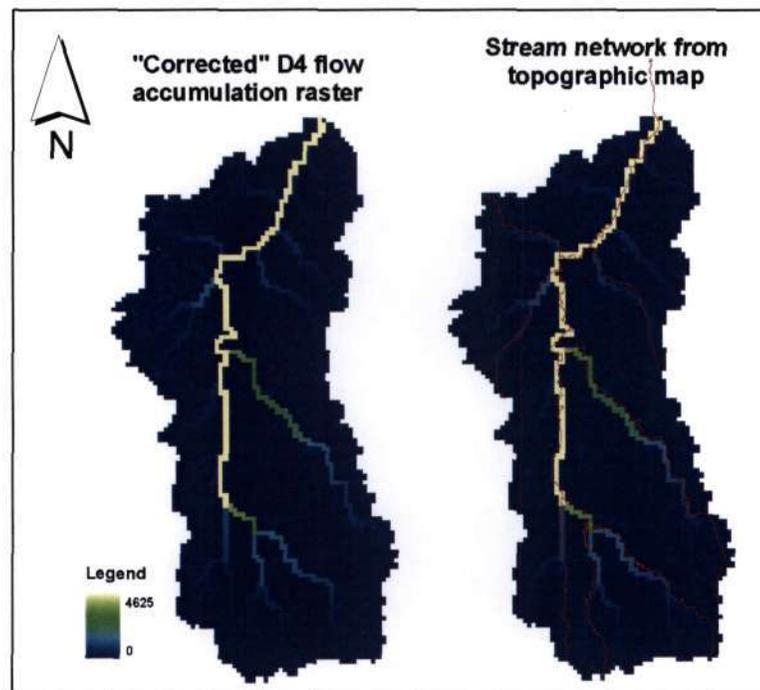


Figure 7-11. The “corrected” flow accumulation raster is shown on the left which is overlain on the right with a stream network (in red) traced from a topographic map (from Midgley et al., 1994) at a spatial scale of 1:250 000.

(ii) Stream networks

The delineation of a catchment's stream network is important in catchment modelling as it determines which cells contain a channel and which do not. The pixels or grid cells of the catchment would then be modelled accordingly. In the TOPKAPI model, the surface flow of a cell that does not consist of a channel will be modelled with an overland store only. For a cell that does contain a channel, the surface flow will be modelled with both an overland store and a channel store and the flow within that cell would be partitioned between the two stores respectively (in the manner explained in Chapter 6).

The stream network delineation is made from the flow accumulation raster by assigning a threshold value to define the minimum number of upslope cells (or area) which are required to initiate a channel. Fig. 7-12 shows the difference when stream networks were delineated using five different threshold values, i.e. 5, 10, 25, 50 and 65 pixels. Since the pixel resolution of the flow accumulation raster used was 1km^2 (from Fig. 7-11), these threshold values are equivalent to areas of 5, 10, 25, 50 and 65km^2 respectively. The drainage densities of the extracted networks are shown in Table 7-1; drainage density (in km^{-1}) is defined as the ratio of total stream length to total catchment area. Since the drainage direction of the flow accumulation raster is based on a D4 method, the total length of the stream network drained is equal to the number of channel pixels (based on the raster's resolution of 1km^2). This can be understood from the fact that the north/south or east/west distance across a cell is 1km given a raster resolution of 1km^2 . This is shown in the third column in Table 7-1. The drainage density (column 4 in Table 7-1) is computed by dividing the total length of the stream network by the total catchment area, which is 4625 km^2 (from 4625 pixels in Fig. 7-11).

Table 7-1. The drainage densities of the extracted stream networks shown in Fig. 7-12.

	Threshold Area (km^2)	No. of channel Pixels (also <i>total stream length</i> in km)	Drainage Density (km^{-1})
Stream net. 5	5	1369	0.296
Stream net. 10	10	1061	0.229
Stream net. 25	25	712	0.154
Stream net. 50	50	539	0.117
Stream net. 65	65	455	0.0984

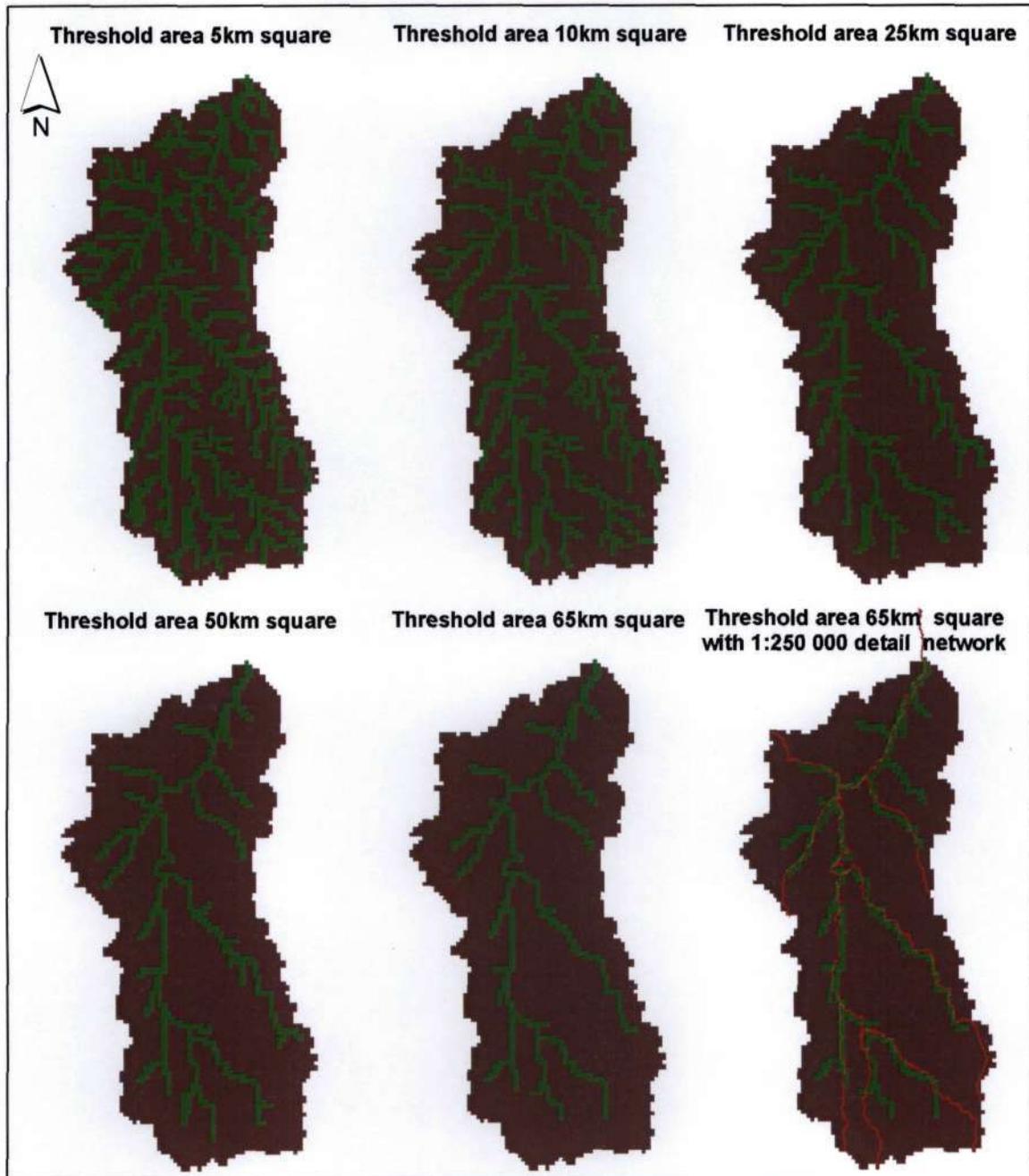


Figure 7-12. The stream networks delineated utilizing five different threshold areas, i.e. 5, 10, 25, 50 and 65km square, on ARCGIS™. The superimposition of the 1:250 000 network (from Midgley et al., 1994) on the 65km square network is shown on the bottom right.

Tarboton et al. (1991) state that in “extracting channel networks from digital elevation models, it is important that the networks extracted be close to what traditional workers using maps or fieldwork would regard as channel networks”. In their paper, Tarboton et al. present a rigorous method for the extraction of stream networks. They base their methods on the morphometric scaling properties of stream networks as discovered by Horton (1932, 1945), Strahler (1952, 1964) and others since. In particular, they make

use of two properties, the constant drop property and the power law scaling property of slope with area, and suggest that the smallest threshold area that should be used be that area for which these scaling properties are still valid.

From the literature on the TOPKAPI model, it is not clear on how a threshold value is chosen. Therefore, in order to simplify matters, the network extracted using a threshold area of 25km² was arbitrarily taken as representative of the catchment's actual network. According to Todini (2005), stream networks in reality only form about 10% to 15% of the catchments' area and that the level of accuracy for this detail is not critical in the TOPKAPI model. The use of a 25km² threshold area gives a drainage density of approximately 15%. This is shown in Fig. 7-12.

A further topographically derived input of the TOPKAPI model is the order of the stream network ordered according to the method of Strahler (1964). The Strahler basin order is defined in Fig. 3-5 in Section 3.2.3 and also in Pegram and Parak (2004) attached in Appendix A. This input, the Strahler basin order, is required to facilitate the estimation of a Manning's channel roughness coefficient n_c for channel stores of the model. Values of n_c should firstly be estimated from a priori knowledge (through fieldwork) of the channel reaches in the catchment, using literature such as Chow (1959) or Barnes (1967) for its estimation. If this is not known, the roughness coefficient of the channel reaches in each pixel can be estimated based on the channel order assigned to each reach using the ordering method of Strahler (1964). In Liu and Todini (2002), channel orders of 1, 2, 3 and 4 were assigned n_c values of 0.045, 0.040, 0.035 and 0.035 m^{-1/3}s respectively for the Upper Reno catchment in Italy. A stream order raster can be computed from the delineated stream network using ARCGIS™. Fig. 7-13 shows the stream orders of the Liebenbergsvlei catchment extracted using a threshold area of 25km².

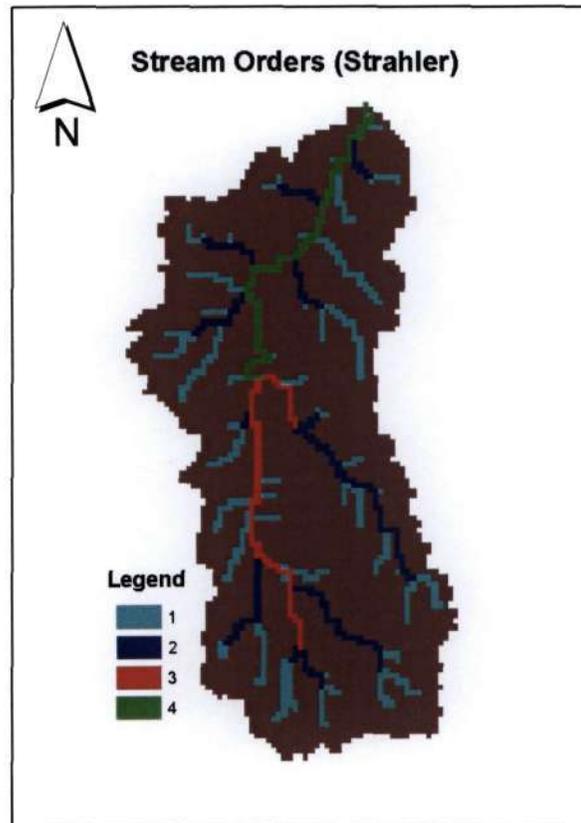


Figure 7-13. The stream orders, ordered according to the method of Strahler (1964), based on a stream network extracted using a threshold area of 25km² on ARCGIS™.

(iii) Surface slopes

A surface slope raster was computed by means of an inbuilt tool on ARCGIS™ using the 1km DEM as input. The output of this function gives the surface slopes of each cell, in degrees or as percentages. The surface slope raster for the Liebenbergsvlei catchment is shown below in Fig. 7-14 where the slopes are shown in degrees and the resolution of the raster is 1km square.

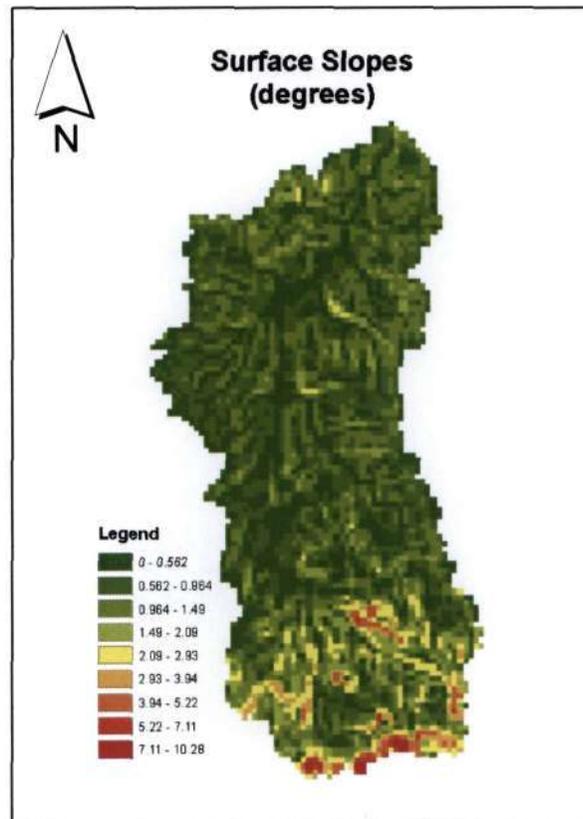


Figure 7-14. A surface slope raster (in degrees) is shown computed from the 1km DEM on ARCGIS™. The resolution of the raster is 1km square.

7.1.3. Soils map

Fig. 7-15 shows a vector-based soils map of South Africa (actually a landtype map with related soils properties) obtained from SIRI (1987). The different soils attributes of the map are represented by the different polygons, where the accuracy of the map (scale of detail) is 2,5km. In order for this map to be of use in this study, it was necessary to be able to identify the soil properties of each pixel. In order to accomplish this, the vector-based soils map of South Africa was converted into raster form using a feature on ARCGIS™, called "Feature to Raster". A cell size of 1km square was specified so that the raster-based soils map would be compatible with the other input rasters. It is important to note that although the raster now has a resolution of 1km square, its accuracy is still at the scale at which it was mapped, i.e. 2.5km. A mask of the Liebenbergsvlei catchment was then used to clip the raster-based soils map for the area of interest. These are shown in Fig. 7-15.

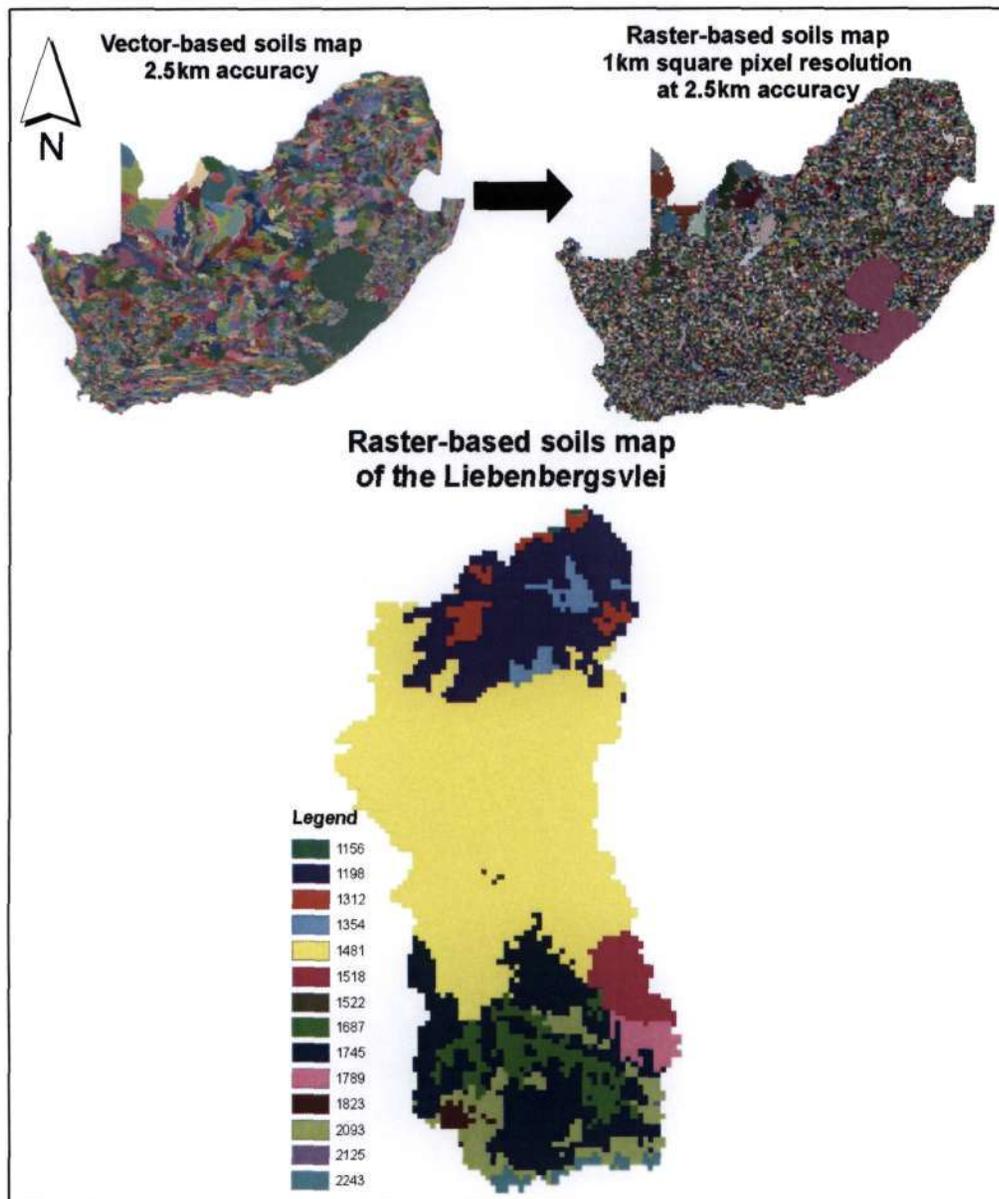


Figure 7-15. A vector-based soils map from SIRI (1987) was converted into a raster with a pixel resolution of 1km square from which the raster-based soils map of the Liebenbergsvlei catchment was clipped. The legend distinguishes soils groups of which the properties are given in Table 7-2.

The legend of the soils map for the Liebenbergsvlei catchment identifies and displays a particular code which has related soil properties, rather than actually identifying the soil type. The related soil properties are tabulated in Table 7-2 and correspond to the codes shown in the legend. The soil properties given there are inter alia, for both an upper top soil layer and a lower sub-soil layer, depth of layer, wilting point, field capacity, porosity and the saturated drainage rates.

Table 7-2. The related soil properties of the landtype map (SIRI, 1987) of the Liebenbergsvlei catchment (see Fig. 7-15).

Land-type code	Depth of top soil (Layer A) (m)	Depth of sub soil (Layer B) (m)	Wilting point (A) (m/m)	Wilting point (B) (m/m)	Field capacity (A) (m/m)	Field capacity (B) (m/m)	Porosity (A) (m/m)	Porosity (B) (m/m)	Saturated drainage rate (A to B) (fraction)	Saturated drainage rate (B to C) (fraction)	Impervious areas next to stream (fraction of landtype)	Other impervious areas (fraction of landtype)	USLE erodibility factor (A) (fraction)	USLE erodibility factor (B) (fraction)	Plant available water (cm)
1156	0.23	0.29	0.23	0.25	0.28	0.30	0.44	0.44	0.21	0.21	0.002	0.040	0.42	0.42	27.9
1198	0.26	0.41	0.22	0.24	0.28	0.30	0.44	0.43	0.21	0.21	0.002	0.050	0.41	0.41	37.7
1312	0.23	0.37	0.18	0.26	0.25	0.31	0.41	0.42	0.24	0.24	0.002	0.003	0.37	0.37	34.8
1354	0.22	0.32	0.16	0.22	0.23	0.29	0.43	0.43	0.20	0.20	0.000	0.008	0.58	0.58	36.3
1481	0.24	0.41	0.14	0.20	0.22	0.27	0.44	0.42	0.24	0.24	0.002	0.006	0.48	0.48	47.9
1518	0.25	0.43	0.15	0.21	0.23	0.28	0.42	0.41	0.31	0.31	0.002	0.040	0.39	0.39	48.7
1522	0.21	0.23	0.15	0.16	0.23	0.24	0.44	0.41	0.22	0.22	0.000	0.140	0.41	0.41	34.8
1687	0.21	0.21	0.18	0.26	0.25	0.31	0.41	0.41	0.24	0.24	0.000	0.005	0.46	0.46	25.0
1745	0.28	0.51	0.13	0.19	0.22	0.27	0.44	0.42	0.32	0.32	0.002	0.010	0.45	0.45	59.9
1789	0.23	0.34	0.18	0.25	0.25	0.30	0.41	0.41	0.29	0.29	0.002	0.010	0.39	0.39	34.7
1823	0.28	0.53	0.14	0.20	0.22	0.27	0.43	0.40	0.44	0.44	0.000	0.040	0.37	0.37	57.9
2093	0.23	0.34	0.13	0.15	0.21	0.23	0.45	0.43	0.39	0.39	0.003	0.240	0.45	0.45	48.0
2125	0.22	0.43	0.18	0.19	0.25	0.26	0.43	0.41	0.44	0.44	0.002	0.080	0.35	0.35	46.4
2243	0.20	0.13	0.23	0.23	0.28	0.29	0.41	0.42	0.25	0.25	0.000	0.610	0.27	0.27	18.3

The inputs required by the TOPKAPI model with regard to the soil store are the depth of the surface soil layer, saturated hydraulic conductivity, the saturated moisture content and the residual moisture content. Although the soils' parameters related to the map do not explicitly give the latter three input parameters required by the model, these may be inferred from the information given. Ideally, a soil types map would be able to identify the particular soil that is found in each pixel. Based on this identification, characteristic soils properties may be found from literature. However, the map used in this research is a landtype map (as already mentioned) with related soil properties. It is included in this dissertation as it was the only map available at no cost, and for informative purposes. However, it is felt that a map which identifies the soil type of each pixel might be better suited for the application of the model.

7.1.4. Landuse map

Landuse parameters are required so that Manning's roughness coefficients may be inferred for the hillslope surfaces of a catchment. Fig. 7-16 shows a raster-based image of the landuse for each pixel of the continent of Africa at a resolution of 1km square (GLCC, 1997), from which a map of the Liebenbergsvlei catchment may be masked. This database is one part of a suite of global land cover characteristics (at a resolution of 1km) for all the continents of the world which are freely downloadable from the United States Geological Survey (USGS) website (<http://edcdaac.usgs.gov/glcc>). The coverage characteristics represent averages thereby giving flexibility with regard to seasonal changes in land use.

The legend of the map identifies the landuse type of each pixel, which in turn can be used to infer Manning's roughness coefficient for the hillslopes of a catchment using literature such as Chow (1959: 108). This input is required for the overland store of the TOPKAPI model. The landuse map for the Liebenbergsvlei catchments shows that the catchment's landuse consists predominantly of cropland and grassland. From Chow (1959: 108), typical values of n_o (the surface roughness coefficient) for this coverage type range from 0.020 to 0.050m^{-1/3}s.

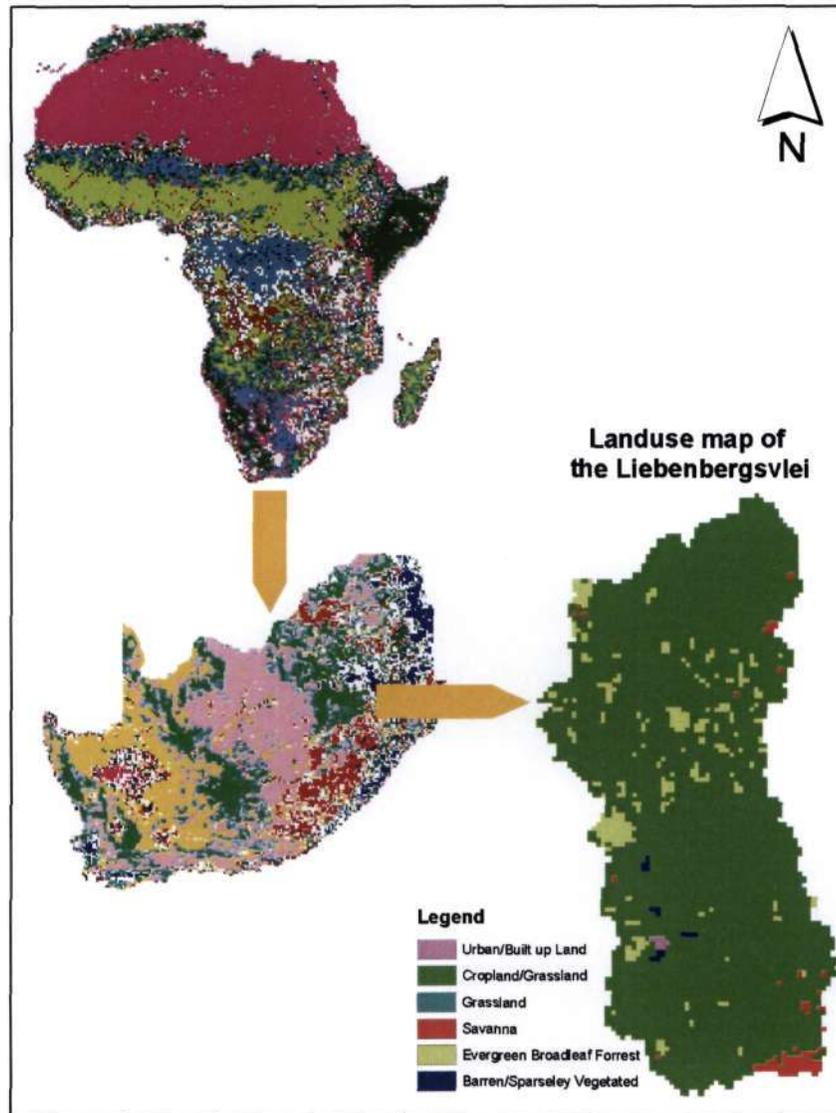


Figure 7-16. A landuse map of Africa, South Africa and the Liebenbergsvlei catchment, all at a pixel resolution of 1km square (GLCC, 1997).

7.1.5. Rainfall input

The TOPKAPI model requires distributed rainfall information in real-time for flood-forecasting purposes. The use of remote sensing techniques, such as satellite and radar estimates of distributed precipitation information, is ideal for this application since they provide precipitation estimates in fine spatial detail over a large area and the pixel format of these precipitation estimates is well-matched with the processing grid cells of the model. However there are errors associated with remotely sensed precipitation information and as such raingauge estimates are used, in combination with the satellite and radar estimates, to condition a “best” merged estimate of real-time distributed precipitation. An example of such a combination technique employed locally is the SIMAR (Spatial Interpolation and Mapping of Rainfall) project which was jointly

undertaken between the South African Weather Service (SAWS) and the University of KwaZulu-Natal (UKZN) under a contract with the Water Research Commission (WRC) (Pegram, 2004). The merging process will not be covered here, but as a point of example, the attainment of real-time distributed radar estimates of rainfall for input into the model will be explained below (using rainfield data from the S-Band MRL5 weather radar covering the Liebenbergsvlei catchment).

Fig. 7-17 shows an instantaneous volume scan of radar reflectivity (dBz) at 2km altitude above ground level (a.g.l.), from which the rainfall for the Liebenbergsvlei catchment has been clipped out. The resolution of the rainfall estimate is 1km². In order for these images to be used as input for the TOPKAPI model, it is worth knowing the rainfall at ground level that occurs over a finite time step (for example an hour). In order to accomplish this, the radar images are kriged down to ground level from all 18 levels a.g.l. (at 1km vertical spacing) and accumulated over an hour (Pegram et al., 2005). Furthermore, the reflectivity values (in dBz) are converted into rainfall intensities (in mm.hr⁻¹) using the Marshall Palmer formula (Marshall and Palmer, 1948). The format of these rainfall estimates make it possible to input the rainfall intensities incident on each grid cell in each time step when modelling each pixel.

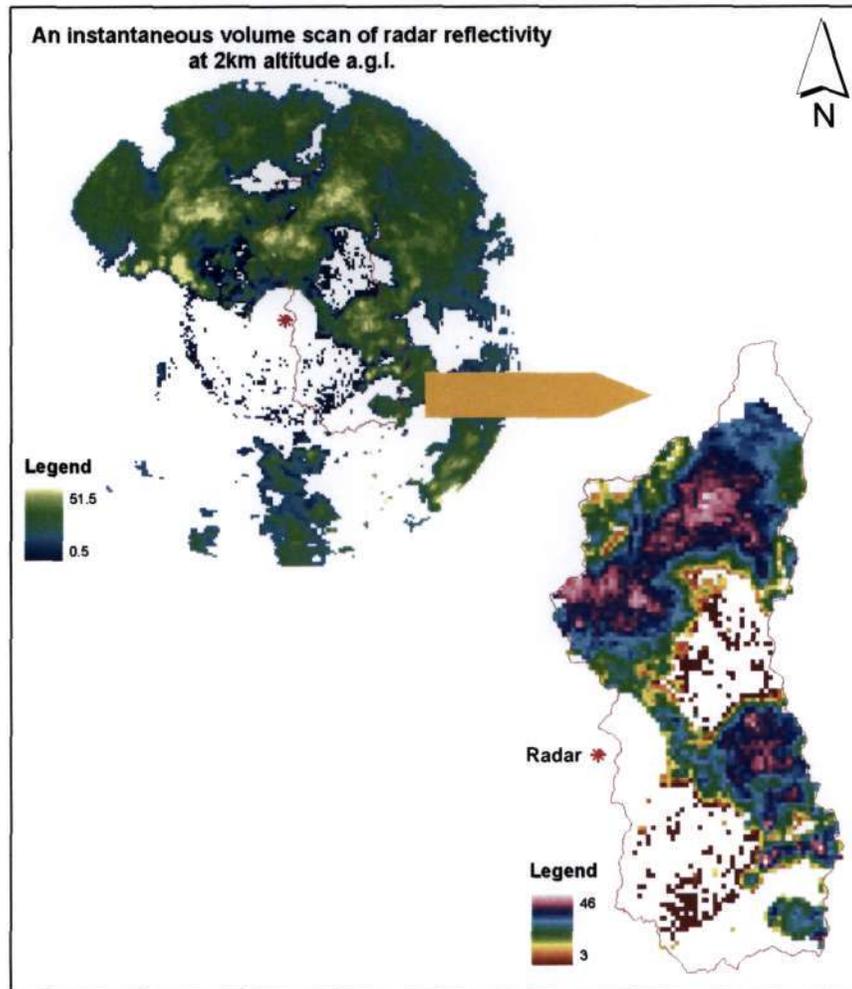


Figure 7-17. An instantaneous volume scan of radar reflectivity (dBz) at 2km altitude above ground level (a.g.l.), from which the rainfall for the Liebenbergsvlei catchment has been clipped out. The resolution of the images is 1km square.

The volume scan from the radar comprises a 200 by 200 matrix at a resolution of 1km square, where the image origin is located at the radar centre (in Cartesian coordinates). In order for the rainfall image to be correctly aligned with the Liebenbergsvlei catchment on ARCGIS™, which take its origins at the lower left corner, the radar image was given a new origin at the lower left corner by shifting it up and to the right by a distance equal to the radius of the image, i.e. 100km (100 000m in Cartesian coordinates). The rainfall estimate for the Liebenbergsvlei catchment could then be clipped out using a mask of the catchment.

7.1.6. Data alignment

In any work that involves geographically referenced data, it is important that the coordinate systems used are the same, especially when one uses data from different sources. Furthermore, it is important that the cells of the data are correctly aligned so that the input data match the correct cell which is being modelled. All GIS work carried out in this research made use of the ARCGIS™ software. A brief explanation of how the data was managed and aligned follows.

Firstly, the combination of data from different sources involved the determination of the geographic coordinate system on which they were based. Secondly, it was necessary to establish whether the data was projected or not. A geographic coordinate system (GCS) uses a three-dimensional spherical surface to define locations on the earth. Points on a GCS are referenced by their longitude (which run north - south) and latitude (which run east - west around the earth) and are measured in degrees from the earth's centre. A projected coordinate system is defined on a flat two-dimensional surface where locations are identified by x-y coordinates on a grid. A projected coordinate system is always based on a GCS which has been converted (projected) using some method. The GCS adopted for mapping in South Africa at present is the WGS84 global ellipsoid while the projection system used (to cover a limited area of 2° longitude) is a Transverse Mercator Map. Fig. 7-18 below shows a screen capture of the geographic and projected coordinate system details, from ARCGIS™, which has been used uniformly in this research to represent all data covering the extent of the Liebenbergsvlei catchment.

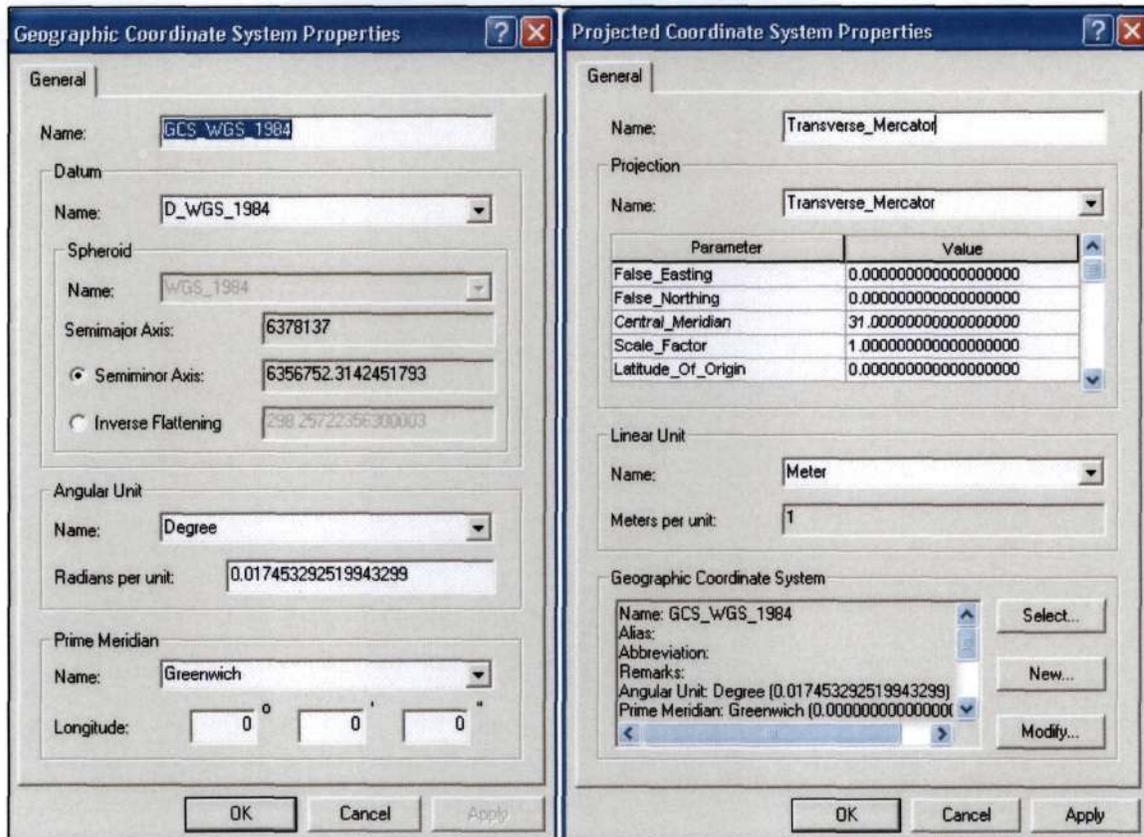


Figure 7-18. A screen capture of the geographic and projected coordinate system details, from ARCGIS™, which has been used uniformly in this research to represent all data covering the extent of the Liebenbergsvlei catchment.

Furthermore, in order to align the data “pixel to pixel”, a mask of the Liebenbergsvlei catchment representing a matrix of 62 columns and 121 rows at a resolution of 1km square was used to clip and extract all data. The mask was created from the DEM of the catchment and ensures that the clipped data, provided it is at the same resolution, has the same number of columns and rows and is originated at the same point (the lower left corner in ARCGIS™). Fig. 7-19 below shows a screen capture of all the required input data in text-format. These data can be displayed in ARCGIS™ in raster-format, but are needed in text-format for input into a code to run the TOPKAPI model. It can be seen from Fig. 7-19 that all the data have the same origin, number of columns and rows and the same cell size, all of which would ensure “pixel to pixel” alignment.

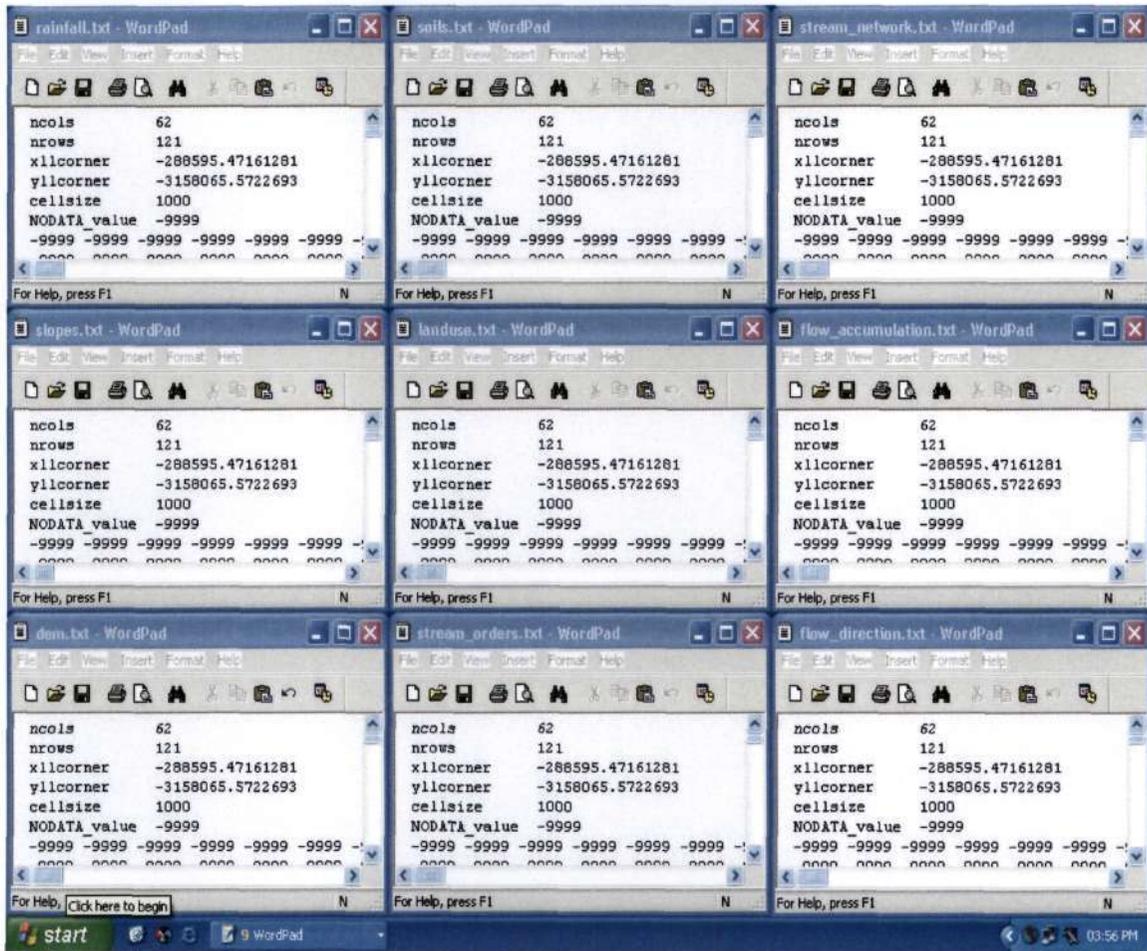


Figure 7-19. A screen capture of the headings of the files of all the input data, in text-format, which describe the properties of each pixel of the Liebenbergsvlei catchment, which is required by the TOPKAPI model. These data can be displayed in ARCGIS™ in raster-format. The figure shows that all the data consist of 62 columns, 121 rows, are aligned at the lower left corner in a projected coordinate system (see Fig. 7-18) and are at a cell size resolution of 1000m square.

7.2. Test application

The test application of the TOPKAPI model consisted of creating a four-cell generic catchment, together with establishing the intricate intra- and inter-cell operations. This test was simply created on a standard spreadsheet package using Microsoft Excel™. The purpose of this small-scale application of the TOPKAPI model was to test our understanding of the model and form the basis for modelling an entire catchment using a higher level of programming language such as C++. This latter task is not accomplished in this research and is left for completion in a follow up study. Instead, the aim of this research was to lay the groundwork in preparation for the establishment

of the TOPKAPI model as a fully-functioning real-time rainfall-runoff application for flood-forecasting purposes in the Liebenbergsvlei catchment and others.

The four cell generic catchment was imagined as follows: Cell 1 flows into Cell 2, which in turn flows into Cell 3, which in turn flows into Cell 4. Channel flow was only initiated in Cell 2 and thus Cells 3 and 4 had channel flow as well. In order to create the intra- and inter-cell operations of the test catchment, it was necessary to model the three fundamental components of the TOPKAPI model in each cell, i.e. the soil, overland and channel (except for Cell 1) stores respectively. The model was then run based on pulsed precipitation inputs (in each time step) and a simple continuity check was used to verify the operations of the test catchment. An explanation of how this was accomplished is detailed in the sub-sections that follow.

7.2.1. Input parameters

The initial values for the input parameters were arbitrarily chosen but were kept within the range of expected values for the given variables, which were suggested by Liu and Todini (2002). The initial chosen parameters for each cell, together with their suggested range (shown in parentheses in the second column) are given below in Table 7-3. The parameters that would remain unchanged for all the cells of a catchment are the horizontal dimensions of the grid cell, taken as 1000m, the time step, taken as 1 hour (3600s), the non-linear exponents α , which was taken as 3 for the soil stores and $5/3$ (from Manning's formula) for the overland and channel stores respectively, and the parameters related to the computation of the channel width (from Eq. 6-24 in Chapter 6).

Table 7-3. Input parameters chosen for the “four cell generic catchment” for the test application of the TOPKAPI model on Microsoft Excel™. The suggested range of the parameters (in Liu and Todini, 2002) is given in parentheses in the second column.

Parameters		Cell 1	Cell 2	Cell 3	Cell 4
Soil Store					
Depth of surface soil layer (m)	L (0.1–2)	0.5	0.75	1.00	1.25
Saturated hydraulic conductivity (m.s ⁻¹)	k _s (10 ⁻⁶ –10 ⁻³)	0.001	0.001	0.001	0.001
Surface Slope	tanβ	0.09	0.08	0.07	0.05
Residual soil moisture content	θ _r (0.01–0.1)	0.04	0.05	0.06	0.07
Saturated soil moisture content	θ _s (0.25–0.7)	0.45	0.50	0.55	0.60
Overland Store					
Manning's surface roughness coeff.	n _o (0.05–0.4)	0.1	0.2	0.3	0.4
Surface slope	tanβ	same as for soil store			
Channel Store					
Manning's channel roughness coeff.	n _c (0.02–0.08)	0.03	0.04	0.05	0.06
Surface slope	tanβ	same as for soil store			
Constant Parameters for all cells					
Horizontal dimension of cell (m)	X	1000			
Time step (s)	Δt	3600			
Non-linear soil exponent	α _s (2–4)	3			
Non-linear overland exponent	α _o	5/3			
Non-linear channel exponent	α _c	5/3			
Max. channel width at outlet (m)	W _{max}	10			
Min. channel width for A _{threshold} (m)	W _{min}	1			
Area required to initiate channel (m ²)	A _{threshold}	1 000 000			
Total area drained by catchment (m ²)	A _{total}	4 000 000			
Area drained by <i>i</i> th cell (m ²)	A _{drained}	1x10 ⁶	2x10 ⁶	3x10 ⁶	4x10 ⁶

7.2.2. Soil store

In Chapter 6 it was explained that the soil store of the TOPKAPI model is the regulating store of each cell. The operations of the soil store for the “four cell generic catchment” were formulated on Microsoft Excel™ as shown in Fig. 7-20. An explanation of how some of the key functions of the soil store for Cell 1 operate in each column is listed after Fig. 7-20.

Cell 1: Soil Store

L	0.5 m
k_s	0.001 m.s ⁻¹
tan β	0.09
θ_i	0.04
θ_s (porosity)	0.45
V_{sum}	205000 m ³
C_s	0.005223 m ² /s
$C_s X X^{2n}$	5.223E-18 m ⁷ /s

1	2	3		4	5	6	7	8
Time Step	Begin of Time Step t_n (s)	Precipitation Input		Upstream Inflow $(1-\lambda)W(X)Q^u$ (m ³ s ⁻¹)	Total Input (m ³ s ⁻¹)	Moisture Stored @ t_n $V_s(t_n)$ (m ³)	Outflow @ t_n $Q_s(t_n) = (C_s X X^{2n})V^u$ (m ³ s ⁻¹)	
		p (mm/hr)	pX^2 (m ³ s ⁻¹)					
1	0	20	5.556		5.556	5000	0.00000	
2	3600	20	5.556		5.556	15000	0.00002	
3	7200	20	5.556		5.556	28723	0.00012	
4	10800	20	5.556		5.556	44428	0.00046	
5	14400	20	5.556		5.556	61242	0.00120	
6	18000	20	5.556		5.556	78736	0.00255	
7	21600	20	5.556		5.556	96677	0.00472	
8	25200	20	5.556		5.556	114929	0.00793	
9	28800	20	5.556		5.556	133405	0.01240	
10	32400	20	5.556		5.556	152043	0.01836	

9												10												11												12												13												14												15												16												17												18												19												20											
Quasi-analytical solution to the non-linear differential reservoir equation for the soil store																																																																																																																																															
$u=y^{(n-1)}$	$u^{(n-1)}$	Fitted parameters		$y^{(n-1)}y^{(n-1)}+\alpha y=0$	A	B	C	P_1	P_2	$u(t+\Delta t)$	$y(t+\Delta t)$																																																																																																																																				
4.000E-08	8.000E-12	5000	0	0	-55556	0	-1.880E-22	1.371E-11	-1.371E-11	4.444E-09	15000																																																																																																																																				
4.444E-09	2.963E-13	15000	0	0	-166666	0	-6.268E-23	7.917E-12	-7.917E-12	1.212E-09	28723																																																																																																																																				
1.212E-09	4.220E-14	28723	0	0	-319140	0	-3.273E-23	5.721E-12	-5.721E-12	5.066E-10	44428																																																																																																																																				
5.066E-10	1.140E-14	44428	0	0	-493640	0	-2.116E-23	4.600E-12	-4.600E-12	2.866E-10	61242																																																																																																																																				
2.866E-10	4.354E-15	61242	0	0	-680468	0	-1.535E-23	3.918E-12	-3.918E-12	1.613E-10	78736																																																																																																																																				
1.613E-10	2.049E-15	78736	0	0	-874843	0	-1.194E-23	3.456E-12	-3.456E-12	1.070E-10	96677																																																																																																																																				
1.070E-10	1.107E-15	96677	0	0	-1074190	0	-9.729E-24	3.119E-12	-3.119E-12	7.571E-11	114929																																																																																																																																				
7.571E-11	6.587E-16	114929	0	0	-1276991	0	-8.181E-24	2.860E-12	-2.860E-12	5.619E-11	133405																																																																																																																																				
5.619E-11	4.212E-16	133405	-8.77449E-22	0	-1482273	6.57735E-27	-7.048E-24	2.655E-12	-2.655E-12	4.326E-11	152043																																																																																																																																				
4.326E-11	2.845E-16	152043	0	0	-1689369	0	-6.184E-24	2.487E-12	-2.487E-12	3.428E-11	170801																																																																																																																																				

21	22	23	24	25	26	27
End of Time Step $t_n=t_n+\Delta t$ (s)	Intermediate Moisture St. $V_s(t_n)$ (m ³)	Saturation Excess e_s (m ³)	Evapotrans Losses $E_s X^2$ (m ³)	Actual Moisture St. $V_s(t_n)$ (m ³)	Outflow @ t_n $Q_s(t_n) = (C_s X X^{2n})V^u$ (m ³ s ⁻¹)	Avg Outflow over Δt $Q_s(\Delta t)$ (m ³ s ⁻¹)
3600	15000	0		15000	0.00002	9.141E-06
7200	28723	0		28723	0.00012	7.070E-05
10800	44428	0		44428	0.00046	2.909E-04
14400	61242	0		61242	0.00120	8.289E-04
18000	78736	0		78736	0.00255	1.875E-03
21600	96677	0		96677	0.00472	3.635E-03
25200	114929	0		114929	0.00793	6.325E-03
28800	133405	0		133405	0.01240	1.017E-02
32400	152043	0		152043	0.01836	1.538E-02
36000	170801	0		170801	0.02603	2.219E-02

Figure 7-20. The operations of the soil store of Cell 1 for the first ten time steps as modelled using Microsoft Excel™.

- **Col. 1:** 75 time steps were created for this model which corresponds to a total storm duration of 75 hours (approximately 3 days).
- **Col. 6:** The total input in each time step is taken as the sum of the incident precipitation (lumped over the time step) and any contributions from the soil and overland stores of an upstream cell. This latter input is derived from the average outflow over the time step from the upstream cell. In this instance, Cell 1 is a source cell with no upstream contributors (and hence Col. 5 is blank), however it is worth noting the partitioning performed with regard to this input, i.e. the

partitioning of flow to the channel of a channel cell and to the soil store of the next downstream cell. This is explained in Section 7.2.4.

- **Col 7:** The initial soil moisture stored was set at 5000m^3 , which forms approximately 2.4% of the saturated moisture volume V_{sm} (which is $205\,000\text{m}^3$ for Cell 1). In Liu et al. (2005), the initial soil saturation percentage was set at the same value (0.9%) for all the cells of the Upper Xixian catchment (in China) when the calibration for the catchment was performed using the TOPKAPI model.

- **Col 8:** The sub-surface outflow at the beginning of the time interval is a function of the volume stored at the beginning of the time interval, which is computed from the outflow term of Eq. 6-14 (in Chapter 6), i.e.

$$Q_{s_i}(t_0) = \frac{C_{s_i} X}{X^{2\alpha_{s_i}}} V_{s_i}^{\alpha_{s_i}} \quad [\text{m}^3\text{s}^{-1}].$$

- **Col. 9-20:** These columns are necessary to compute the quasi-analytical solution offered by Liu and Todini (2002) shown in Section 6.1.6 of Chapter 6, where y represents the volume term V , c the non-linear exponent α , u a substitution variable used for the integration and β and α (not to be confused with α from the non-linear exponent) are two variables fitted so as to approximate the non-linear term y^c with a second order polynomial $\beta y^2 + \alpha y$. The variables β and α were fitted by solving the equation $y^c - \beta y^2 + \alpha y = 0$ by

$$\text{iterating } \beta \text{ and } \alpha \text{ such that } \beta = \frac{y^c + \alpha y}{y^2} \text{ and } \alpha = \frac{\beta y^2 - y^c}{y}.$$

The solution of this equation is shown in Col. 13 and the variables were fitted in each time step by activating the *iteration* function on Microsoft Excel™. However, as explained in Section 6.1.6 of Chapter 6, it turns out that the best fit of these variables is $\beta = y^{c-2}$ and $\alpha = 0$. Cols. 14-18 are used as intermediary steps to compute u (Col. 19) at the end of the time step, i.e. at $t_1 = t_0 + \Delta t$. This parameter (u) is then back substituted to obtain, in Col. 20, $y(t_0 + \Delta t)$ which is the solution of the non-linear soil reservoir equation.

- **Col. 22:** The intermediate moisture storage $V'_s(t_1)$ is the solution of the non-linear reservoir equation for the soil store and is equal to Col. 20 when input is greater than zero. When input equals zero while there is moisture stored in the soil store at the beginning of the time interval, the non-linear differential reservoir equation for the soil store reduces to a decay function and the intermediate moisture storage at the end of the time is computed from Eq. 6-45 in Chapter 6. When input and initial moisture storage, $V_s(t_0)$, equals zero at the

beginning of the time interval, $V'_s(t_i) = 0$. These switches are achieved using two *if statements* imbedded in each cell of this column on Microsoft Excel™.

- **Col. 23:** Saturation excess is the input to the overland store and is activated upon the saturation of the soil store, i.e. when $V'_s(t_i) \geq V_{sm}$. It is taken as an average excess given off during the time step and becomes the input for the overland store of that time step. The switch for this is also achieved using an *if statement*.
- **Col. 24:** Evapotranspiration losses are subtracted as a lumped loss at the end of the time step from the intermediate soil storage. In this exercise, a suitable method to compute this amount has not been implemented (see Section 6.1.4 in Chapter 6) and hence the column is blank.
- **Col. 25:** The actual moisture storage at the end of the time step $V_s(t_i)$ results from the subtraction of the saturation excess and evapotranspiration losses from the intermediate moisture storage $V'_s(t_i)$.
- **Col. 26:** The sub-surface outflow at the end of the time interval is a function of the volume stored at the end of the time interval, which is computed from the outflow term of Eq. 6-14 (in Chapter 6), i.e. $Q_{s_1}(t_i) = \frac{C_{s_1} X}{X^{2\alpha_s}} V_{s_1}^{\alpha_s} \quad [m^3s^{-1}]$.
- **Col. 27:** The average sub-surface outflow over the time interval Δt is computed simply from the average of the outflow at the beginning of the time interval and the end of the time interval, i.e. the average of Col. 8 and 26. This then becomes the input, for that time step, for the soil store of the downstream cell. If a channel exists in this cell, then the average sub-surface outflow would need to be partitioned between the channel component of the cell and the soil store of the next downstream cell (see Section 7.2.4).

7.2.3. Overland store

The operations of the overland store for Cell 1 are shown below in Fig. 7-21. The time steps shown are from *time step 1* onwards. However it should be noticed that the overland store is only activated when the soil store becomes saturated during *time step 12*. Initially, it is assumed that there is no water stored on the surface slopes and hence the initial volume is zero. An explanation of how some of the columns operate in modelling the overland store of Cell 1 is listed after Fig. 7-21. Explanations of those

columns that are not given were deemed to be self-explanatory or have been covered in Section 7.2.2.

Cell 1: Overland Store

r_0	0.1
$\tan\beta$	0.09
C_0	3
$C_0 X(t)^{2.5}$	0.0000003

1	2	3	4	5	6	7
Time Step	Begin of Time Step t_0 (secs)	End of Time Step $t_0+\Delta t$ (secs)	Saturation Excess e_0 (m^3)	Overland Input $r_0 X^2 = e_0/\Delta t$ ($m^3 s^{-1}$)	Moisture Stored @ t_0 $V_0(t)$ (m^3)	Outflow @ t_0 $Q_0(t_0) = (C_0 X(t_0)^{2.5}) V_0$ ($m^3 s^{-1}$)
1	0	3600	0	0	0	0.0000
2	3600	7200	0	0	0	0.0000
3	7200	10800	0	0	0	0.0000
4	10800	14400	0	0	0	0.0000
5	14400	18000	0	0	0	0.0000
6	18000	21600	0	0	0	0.0000
7	21600	25200	0	0	0	0.0000
8	25200	28800	0	0	0	0.0000
9	28800	32400	0	0	0	0.0000
10	32400	36000	0	0	0	0.0000
11	36000	39600	0	0	0	0.0000
12	39600	43200	3546	0.984902202	0	0.0000
13	43200	46800	18931	5.258498936	1741	0.0756
14	46800	50400	18931	5.258498936	12930	2.1370
15	50400	54000	18931	5.258498936	18940	4.0374

Quasi-analytical solution to the non-linear differential reservoir equation for the overland store									
y^2	Fitted parameters		$y^2 - y^2 + \alpha y = 0$	A	B	C	p_1	p_2	$Y(t+\Delta t)$
	β	α							
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
0.000E+00	1	0	0	-0.0000003	0	0	0.00	0.00	#DIV/0!
2.520E+05	0.083125877	0	0	-2.49378E-08	0	-210864901.6	14521.19	-14521.19	12930
7.123E+06	0.042605117	0	0	-1.27815E-08	0	-411413723.9	20283.34	-20283.34	18940
1.346E+07	0.037514825	-2.95026E-13	1.11759E-08	-1.12544E-08	-7.86425E-12	-467237408.3	21615.68	-21615.68	21127

18	19	20	21
End of Time Step $t_1 = t_0 + \Delta t$ (secs)	Moisture Storage @ t_1 $V_1(t_1)$ (m^3)	Outflow @ t_1 $Q_1(t_1) = (C_0 X(t_1)^{2.5}) V_1$ ($m^3 s^{-1}$)	Avg. Outflow over Δt $Q_1(\Delta t)$ ($m^3 s^{-1}$)
3600	0	0.0000	0.0000
7200	0	0.0000	0.0000
10800	0	0.0000	0.0000
14400	0	0.0000	0.0000
18000	0	0.0000	0.0000
21600	0	0.0000	0.0000
25200	0	0.0000	0.0000
28800	0	0.0000	0.0000
32400	0	0.0000	0.0000
36000	0	0.0000	0.0000
39600	0	0.0000	0.0000
43200	1741	0.0756	0.0378
46800	12930	2.1370	1.1063
50400	18940	4.0374	3.0872
54000	21127	4.8436	4.4405

Figure 7-21. The operations of the overland store for Cell 1 as modelled on Microsoft Excel™.

- **Col. 4:** This is the input into the overland store and results from the saturation of the soil store of Cell 1. Saturation excess from the soil store is computed at the end of each time step, but it is taken as having being exfiltrated during that time step and so becomes the input for the overland store for that time step.
- **Col. 8-17:** These columns compute the quasi-analytical solution to the non-linear differential reservoir equation for the overland store. The explanations of these are similar to those given in Section 7.2.2 except that in this instance a substitution variable (for the integration) is not used. For certain time steps in Col. 17, results of *#DIV/0!* are returned by Microsoft Excel™ because for these time steps the initial volume stored on the surface and the overland input are zero. This does not effect the final result (as seen in Col. 19) since these are still intermediary steps.
- **Col. 19:** This column records the moisture storage at the end of the time step. If the input into the overland store at the beginning of the time step is greater than zero, the value that Col. 19 takes on is equivalent to Col. 17, which is computed using the quasi-analytical solution calculated in columns 8-17. If input and moisture storage at the beginning of the time interval is zero, then the moisture storage at t_i is zero. However, if the input is zero while there is still moisture stored on the surface at the beginning of the time interval, then the non-linear differential reservoir equation for the overland store reduces to a decay function and the moisture storage at the end of the time interval is computed from Eq. 6-45 in Chapter 6.
- **Col. 21:** The average overland outflow over the time interval Δt is computed simply from the average of the outflow at the beginning of the time interval and the end of the time interval, i.e. the average of Col. 7 and 20. This then becomes the input, for that time step, for the soil store of the downstream cell. If a channel exists in this cell, then the average overland outflow would need to be partitioned along with the sub-surface outflow between the channel component of the cell and the soil store of the next downstream cell (see Section 7.2.4).

7.2.4. Flow partitioning

Flow partitioning is necessary to split the average outflow from the soil and overland store of a cell between the channel of that cell and the soil store of the next cell downstream cell. As explained in Chapter 6 (Section 6.1.3c), this split is proportional to the ratio of the width of the channel of Cell i (W_i) to the overall width of the cell. This is

seen in Section 7.2.2, where the upstream contribution to the soil store of Cell 1 (Col. 5 in Fig. 7-20) has been partitioned using the following proportion $\left(1 - \lambda \frac{W_i}{X}\right)$. The addition of an extra parameter λ was made in order to create a switch which could either activate or deactivate the channel store of a cell and regulate the amount of flow into it (and hence the flow to the downstream cell as well). This switch is explained in Table 7-4 below and the text that follows it.

Table 7-4. The range of values that λ can take when regulating the flow Q to the channel of a cell and to the downstream cell.

λ	Flow to channel $(\lambda \cdot W/X) \cdot Q$	Flow to next cell $(1 - \lambda \cdot W/X) \cdot Q$
0	0	Q
1	$W/X \cdot Q$	$(1 - W/X) \cdot Q$
X/W	Q	0

It is evident from Table 7-4 that the range of λ is from 0 to X/W . If $\lambda=0$, all flow from the soil and overland store of a source cell progress to the soil store of the downstream cell (and hence a channel does not exist for the source cell). If $\lambda=X/W$, all flow from the soil and overland store of a source cell progress to the channel of that cell. This latter scenario is realised at the outlet of a catchment where all the outflows from all the stores of a catchment come together.

It was further felt that this value (λ) could be used to either increase or decrease the amount of flow feeding a channel, since in reality the proportion equivalent to W/X is very small. In Liu et al. (2005), the values of W_{min} and W_{max} that were chosen for the Upper Xixian catchment (with an area of approximately 10 000km²) was 1m and 400m respectively. At the 1km square modelling resolution, this forms a partitioning proportion of approximately 0.1% at the point of channel initiation and 40% at the catchment outlet respectively. However, according to Todini (2005), the proportion of channel cells in a catchment is approximately 10 to 15% of the total number of cells in a catchment. Thus the use of λ as a tool to increase the proportion of flow to a channel does not make a big difference in the overall modelling of the catchment processes and resultant outflow from the catchment. Hence, for this exercise, the use of λ was limited to a switch and would not take on other values besides 0, 1 and X/W .

Fig. 7-22 shows the flow partitioning performed for Cell 1 of the “four cell generic catchment”. In this case, the width of the channel was computed to be 1m (from Eq. 6-24 in Chapter 6, using the parameters given in Table 7-3 above) and λ was set at 0 since Cell 1 does not consist of a channel. The flow that feeds the channel (Col. 8) comprises the sum of the average outflows from the soil and the overland store of that cell (Col. 4 and 5 respectively), which is partitioned in the manner explained above. The remainder feeds the soil store of the downstream cell (Col. 7).

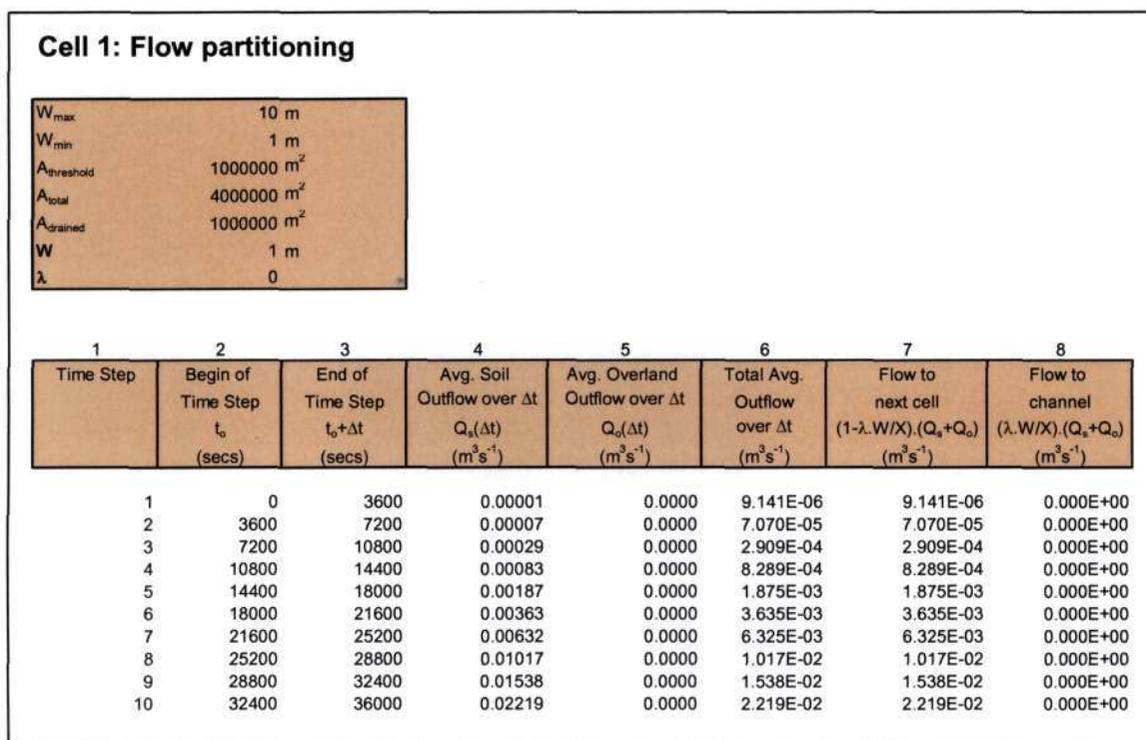


Figure 7-22. The flow partitioning operations of Cell 1, where flow is partitioned between the channel store of Cell 1 and the soil store of the downstream cell, i.e. Cell 2.

7.2.5. Channel store

Fig. 7-23 below shows the channel store for Cell 2 of the “four cell generic catchment”, since Cell 1 did not contain a channel. An explanation of how some of the columns operate in modelling this store of Cell 2 is listed after Fig. 7-23. Explanations of those columns that are not given below were deemed to be self-explanatory or have been covered in the previous sections.

Cell 2: Channel Store						
n_c	0.04					
$\tan\beta$	0.08					
C_c	7.071067812					
$C_c W/(XW)^c$	2.51018E-05					

1	2	3	4	5	6	7
Time Step	Begin of Time Step t_0 (secs)	Input from cell $r_c XW = (\lambda_c W/X)(Q_c + Q_c^u)$ ($m^3 s^{-1}$)	Upstream Channel Input Q_c^u ($m^3 s^{-1}$)	Total Input ($m^3 s^{-1}$)	Moisture Stored @ t_0 $V_c(t_0)$ (m^3)	Outflow @ t_0 $Q_c(t_0) = (C_c W/(XW)^c) V_c^m$ ($m^3 s^{-1}$)
1	0	3.033E-08	0	3.03344E-08	22.35	0.00445
2	3600	1.778E-07	0	1.778E-07	13	0.00181
3	7200	6.253E-07	0	6.25252E-07	9	0.00092
4	10800	1.617E-06	0	1.61715E-06	6	0.00054
5	14400	3.433E-06	0	3.43286E-06	5	0.00034
6	18000	6.372E-06	0	6.37161E-06	4	0.00024
7	21600	1.075E-05	0	1.07453E-05	3	0.00017
8	25200	1.687E-05	0	1.68748E-05	3	0.00013
9	28800	2.509E-05	0	2.50885E-05	2	0.00010
10	32400	3.572E-05	0	3.57213E-05	2	0.00009

Quasi-analytical solution to the non-linear differential equation for the channel store									
y^0	Fitted parameters		$y^c - \beta y^z + \alpha y = 0$	A	B	C	p_1	p_2	$y(t+\Delta t)$
	β	α							
1.774E+02	0.354993004	0	0	-8.91096E-06	0	-0.003404168	0.06	-0.06	13
7.204E+01	0.425091617	0	0	-1.06706E-05	0	-0.016659399	0.13	-0.13	9
3.665E+01	0.486609002	0	0	-1.22148E-05	0	-0.051188217	0.23	-0.23	6
2.139E+01	0.541924701	-1.69628E-15	2.13163E-14	-1.36033E-05	-3.13011E-15	-0.118879594	0.34	-0.34	5
1.370E+01	0.592431093	0	0	-1.48711E-05	0	-0.230841066	0.48	-0.48	4
9.394E+00	0.638890001	0	0	-1.60373E-05	0	-0.397299729	0.63	-0.63	3
6.800E+00	0.681557972	-2.81196E-16	1.77636E-15	-1.71083E-05	-4.12578E-16	-0.628072652	0.79	-0.79	3
5.161E+00	0.720205607	0	0	-1.80785E-05	0	-0.933419833	0.97	-0.97	2
4.100E+00	0.754107981	0	0	-1.89295E-05	0	-1.325365335	1.15	-1.15	2
3.418E+00	0.782081652	0	0	-1.96317E-05	0	-1.81957464	1.35	-1.35	2

18	19	20	21
End of Time Step $t_1 = t_0 + \Delta t$ (secs)	Moisture Storage @ t_1 $V_c(t_1)$ (m^3)	Outflow @ t_1 $Q_c(t_1) = (C_c W/(XW)^c) V_c^m$ ($m^3 s^{-1}$)	Avg. Outflow over Δt $Q_c(\Delta t)$ ($m^3 s^{-1}$)
3600	13	0.00181	0.003130455
7200	9	0.00092	0.001364212
10800	6	0.00054	0.000728539
14400	5	0.00034	0.000440505
18000	4	0.00024	0.000289891
21600	3	0.00017	0.00020325
25200	3	0.00013	0.000150114
28800	2	0.00010	0.000116237
32400	2	0.00009	9.43602E-05
36000	2	0.00008	8.05696E-05

Figure 7-23. The operations of the channel store of Cell 2 as modelled on Microsoft Excel™.

- **Col. 5:** The total input for this store results from the partitioning exercise explained in Section 7.2.4, the result of which is shown in Col. 3, and from an upstream channel outflow (Q_c^u shown in Col. 4). This latter input is the result of the average outflow over Δt from a channel in an upstream cell. Since Cell 1

does not consist of a channel and channel flow is only initiated in Cell 2, this latter input (shown in Col. 4) is zero.

- **Col. 6:** The depth of water in a channel reach is taken to increase linearly with the channel width (Liu et al., 2005). For Cell 2, the channel width (computed from Eq. 6-24 in Chapter 6 using the parameters given in Table 7-3 above) is 4.728m. The depth of water was taken as 0.1% of the channel width which corresponds to an initial volume of 22.35m³.
- **Col. 19:** The moisture stored in the channel at time t_1 is computed using the quasi-analytical solution (shown in columns 8-17) if the input and the moisture stored at the beginning of the time interval is greater than zero. If the input goes to zero while there still remains storage in the channel reach at the beginning of the time interval, then the non-linear differential equation reduces to a decay function for which the solution is given in Eq. 6-45 in Chapter 6. If the input and the moisture stored is zero at the beginning of the time interval, then the moisture stored at the end of the time interval remains zero.
- **Col. 21:** The channel outflow to the channel of the next downstream cell is computed as an average flow over the time interval. This becomes the input for the downstream cell in that interval. This value is computed by simply taking the average of the outflows at time t_0 and time t_1 (i.e. the average of Col. 7 and 20).

7.2.5. Running the model

The cells of the “four cell generic catchment” were run on Microsoft Excel™ using the spreadsheet setup explained above. The complete setup of this “catchment” is shown in Appendix D. The inflow and outflow hydrographs for each store of each cell is shown below in the subsections that follow. In each time step, equal rainfall intensities for each cell were input into the soil store and the outflow from each store of each cell was individually modelled using the rainfall-runoff conversion parameters (for the soil store) and routing parameters (for the overland and channel stores), which are given in Table 7-3. In order to make certain that the correct setup of the model was implemented, a simple check was undertaken. This check was to ensure that *continuity* was maintained for each store and the catchment as a whole, i.e. to ensure for all the cells that the overall input volume minus the output volume matched the volume that remained behind as storage minus the initial storage volumes. Since at this stage of the model's implementation there are no external moisture losses in the form of groundwater recharge or evapotranspiration, the continuity check was easy to quantify. The

outcome of this test highlighted the fact that in the model's original implementation, continuity is not implicitly maintained. The root cause of this anomaly was identified and explained in Section 6.1.6 of Chapter 6 and is explained below. A revised model setup was then implemented to overcome this issue. These exercises are described below and are shown together with all the hydrographs of the test model.

(a) Continuity check

The precipitation input that was used in the "generic catchment's" setup was taken as the same for all the cells and specifically chosen to exercise the model over all plausible ranges of behaviour. This comprised of the following intensities over the 75 time steps: 20mm/hr for the first 25 time steps, thereafter 10mm/hr for the next 15 time steps, thereafter 5mm/hr for the next 20 time steps and finally no input (0mm/hr) for the remaining 15 time steps. Since the input into each cell is taken as constant over the time interval, i.e. a pulsed input, the total volumetric input of the precipitation over all 75 time steps is simply the arithmetic sum of the inputs in each time step. This resulted in a total input volume of 750 000m³ of rainfall into each cell over the whole test period (75 time steps). Based on this input, the intra- and inter-cell outflows were modelled on the spreadsheet for each time step. The outflows from each store of each cell are taken to accrue as averages over the time step, and thus the total volumetric outflow at the end of the test period for each store of each cell was also computed from the sum of the average outflows in each time step. Therefore, continuity was checked by simply verifying that over the 75 time steps, the total input volume to the store minus the total volumetric outflow from the store matched the volume remaining in the store at the end of the final time step minus the initial volume at the beginning of the first time step (see Eq. 7-1). Table 7-5 gives a summary of these flows. Reference should also be made to Appendix D where the complete operations of each store of each cell are shown for all 75 time steps.

The form of the equation used to check continuity in each store of each cell was:

$$\sum Input - \sum Outflow = Vol. Remaining - Initial Vol. \quad (7-1)$$

where $\sum Input$ is the input volume into the store of a cell, $\sum Outflow$ is the outflow volume from a store of a cell, and the *Vol. Remaining* and *Initial Vol.* terms is the volume that is left in the store at the end of the final time step and the volume that is found in the store

at the beginning of the first time step respectively. For the soil store, the $\Sigma Input$ term comprises the sum of precipitation input and any inflow from soil stores of upstream cells. For the overland store, the $\Sigma Input$ term is the saturation excess that is given off from the soil store of that cell. For the channel store, the $\Sigma Input$ term is the sum of the commensurate soil and overland outflows (see Section 6.1.5 in Chapter 6) of that cell and any inflows from channel stores of upstream cells. $\Sigma Outflow$ comprises direct soil drainage and saturation excess for the soil store, direct overland runoff for the overland store and direct channel flow for the channel store. A more convenient form of Eq. 7-1 in order to check if continuity is maintained is:

$$\Sigma Input - \Sigma Outflow - Vol. Remaining + Initial Vol. = Error \quad (7-2)$$

If continuity is preserved, then the *Error* term will equal zero. If continuity is not maintained, then the *Error* term will not equal zero and the magnitude of this term can be used to gauge the degree to which continuity is not upheld. Eq. 7-2 was applied to each store of each cell of the generic catchment and the results are shown in the last column of Table 7-5, with regard to the *Error* observed (or lack of continuity).

The “*Error*” values observed in the last column of Table 7-5 indicate that continuity in each store was not maintained in this test. The root cause of this lack of continuity is in the manner in which the analytical solution is derived and implemented. As explained in Section 6.1.6 of Chapter 6, the analytical solution is made possible by the approximation of the non-linear volume term (y^c) by a second order polynomial ($y(\alpha + \beta y)$). This approximation is made possible by fitting the variables α and β by least squares. It was shown in Chapter 6 that the best fit of these variables is if $\beta = y^{2-c}$ and $\alpha = 0$ for all values of the exponent c . The result of this is that β is implicitly a function of y and is fitted at the beginning of the time step. Now, as y changes during the time step, β is unable to match this change. Consequently, this setup will have the tendency to underestimate the volume at the end of the time step ($y_{t+\Delta t}$) on the rising limb of the store’s hydrograph and overestimate this on the falling limb of the store’s hydrograph. The results of the continuity check in Table 7-5 reflect that continuity is not maintained, as expected. For example, the imbalance (error) that existed in the soil store of Cell 1 after running the model for 75 time steps is calculated as follows:

$$\Sigma \text{Input} - \Sigma \text{Outflow} - \text{Vol. Remaining} + \text{Initial Vol.} = \text{Error}$$

$$(750\,000 + 0) - (489\,096 + 10\,639) - 202\,612 + 5\,000 = 52\,653 \text{ [m}^3\text{]}$$

Table 7-5. A summary of the "continuity check" on each store of each cell for the "four cell generic catchment" modelled on Microsoft Excel™. The values in the last column were computed from Eq. 7-2.

	Input Vol. (m ³)		Output Vol. (m ³)		Store Vol. (m ³)		Error (m ³)
Cell 1							
Soil Store	Rainfall	Upstream	Sat. Excess	Outflow	Remaining	Initial	52 653
	750 000	0	489 096	10 639	202 612	5 000	
Overland Store	489 096		483 544		Remaining	Initial	4 877
					675	0	
Channel Store	Inflow	Upstream	0		Remaining	Initial	0
	0	0			0	0	
<i>Total</i>							57 530
Cell 2							
Soil Store	Rainfall	Upstream	Sat. Excess	Outflow	Remaining	Initial	70 086
	750 000	494 183	831 311	13 386	337 500	8100	
Overland Store	831 311		819 614		Remaining	Initial	9 123
					2 574	0	
Channel Store	Inflow	Upstream	3 939		Remaining	Initial	14
	3 938	0			8	23	
<i>Total</i>							79 223
Cell 3							
Soil Store	Rainfall	Upstream	Sat. Excess	Outflow	Remaining	Initial	75 199
	750 000	829 061	1 011 120	14 502	490 000	11 760	
Overland Store	1 011 120		991 597		Remaining	Initial	12 617
					6 906	0	
Channel Store	Inflow	Upstream	11 586		Remaining	Initial	18
	7 635	3 939			28	58	
<i>Total</i>							87 834
Cell 4							
Soil Store	Rainfall	Upstream	Sat. Excess	Outflow	Remaining	Initial	69 360
	750 000	998 465	1 020 654	11 851	662 500	15 900	
Overland Store	1 020 654		978 227		Remaining	Initial	26 696
					15 731	0	
Channel Store	Inflow	Upstream	21 494		Remaining	Initial	15
	9 901	11 586			78	100	
<i>Total</i>							96 071

Although the error values appear to be large (especially for the soil stores), the values shown in Table 7-5 accrue over 75 time steps. In order to illustrate the error in each time step, the imbalance is easy to quantify when the system is at equilibrium. This condition is found, for the example of the soil store, when this store is saturated. In this instance all incident precipitation to the store becomes saturation excess directly. Thus, for a rainfall intensity of 20mm/hr ($5.556\text{m}^3/\text{s}$), the volume of water that should be exfiltrated to the overland store is $5.556\text{m}^3/\text{s} \times 3600\text{s} = 20002\text{m}^3$. The corresponding amount that is reflected in the soil store of Cell 1 for the 15th time step (shown in Appendix D) is $19\,939\text{m}^3$. Thus it is clear to see that the analytical solution implemented in its original form underestimates the volume by an amount of 63m^3 (in this instance).

This issue was not resolvable without some form of adjustment to the analytical solution. In simple terms, the problem of imbalance is caused due to a "lack of fit" in the variable β . An alternative solution is to "lose" the imbalances that accrue in each time step through evapotranspiration. This latter alternative is not unreasonable, as the estimation of evapotranspiration is relatively not precise and could easily absorb the imbalances that occur in each time step. This alternative is a viable solution in the actual implementation of the TOPKAPI model for a catchment where estimates of evapotranspiration are available. In the test application of the model presented here, an adjustment factor was deemed to be the simplest manner in which to get continuity to balance since evapotranspiration is ignored.

An adjustment factor was implemented for each store of each cell by multiplying the moisture storage at the end of the time step ($y_{t+\Delta t}$) by the factors shown in the last column of Table 7-6. These factors were determined through the *solver tool* on Microsoft Excel™, by multiplying $y_{t+\Delta t}$ with this factor such that the imbalance at the end of 75 time steps is reduced to zero or near zero. It is evident from the factors in Table 7-6 that the adjustments are relatively slight and that the desired result is obtained, i.e. the errors are negligible and continuity is maintained.

Table 7-6. A summary of the “corrected continuity check” on each store of each cell for the “four cell generic catchment” modelled on Microsoft Excel™. An adjustment factor was calibrated for each store of each cell (last column) to get continuity to hold in each store.

	Input Vol. (m ³)		Output Vol. (m ³)		Store Vol. (m ³)		Error (m ³)	Adjustment Factor
Cell 1								
Soil Store	<i>Rainfall</i>	<i>Upstream</i>	<i>Sat. Excess</i>	<i>Outflow</i>	<i>Remaining</i>	<i>Initial</i>	0	1.0045
	750 000	0	541 721	10 667	202 612	5 000		
Over-land Store	541 721		541 032		<i>Remaining</i>	<i>Initial</i>	1	1.0064
					688	0		
Channel Store	<i>Inflow</i>	<i>Upstream</i>	0		<i>Remaining</i>	<i>Initial</i>	0	1
	0	0			0	0		
<i>Total</i>							1	
Cell 2								
Soil Store	<i>Rainfall</i>	<i>Upstream</i>	<i>Sat. Excess</i>	<i>Outflow</i>	<i>Remaining</i>	<i>Initial</i>	0	1.0031
	750 000	551 699	958 867	13 432	337 500	8100		
Over-land Store	958 867		952 177		<i>Remaining</i>	<i>Initial</i>	0	1.0044
					6 690	0		
Channel Store	<i>Inflow</i>	<i>Upstream</i>	4 574		<i>Remaining</i>	<i>Initial</i>	0	1.0017
	4 565	0			14	23		
<i>Total</i>							0	
Cell 3								
Soil Store	<i>Rainfall</i>	<i>Upstream</i>	<i>Sat. Excess</i>	<i>Outflow</i>	<i>Remaining</i>	<i>Initial</i>	1	1.0025
	750 000	961 044	1 218 201	14 602	490 000	11 760		
Over-land Store	1 218 201		1 203 146		<i>Remaining</i>	<i>Initial</i>	-1	1.0093
					15 056	0		
Channel Store	<i>Inflow</i>	<i>Upstream</i>	13 823		<i>Remaining</i>	<i>Initial</i>	1	1.0085
	9 241	4 574			49	58		
<i>Total</i>							1	
Cell 4								
Soil Store	<i>Rainfall</i>	<i>Upstream</i>	<i>Sat. Excess</i>	<i>Outflow</i>	<i>Remaining</i>	<i>Initial</i>	0	1.0019
	750 000	1 208 507	1 299 967	11 940	662 500	15 900		
Over-land Store	1 299 967		1 272 129		<i>Remaining</i>	<i>Initial</i>	-1	1.0039
					27 839	0		
Channel Store	<i>Inflow</i>	<i>Upstream</i>	26 637		<i>Remaining</i>	<i>Initial</i>	1	1.0005
	12 841	13 823			126	100		
<i>Total</i>							0	

It is apparent from Table 7-6 that, for the soil store, these adjustment factors have the effect of transferring the majority of the imbalance (or error) into saturation excess and pushing it through the overland store. This is evidenced from the fact that, in comparing the values for the soil store of Cell 1 from Table 7-5 and 7-6 respectively, the input ($750\,000\text{m}^3$), remaining ($202\,612\text{m}^3$) and initial ($5\,000\text{m}^3$) volumes remain unchanged, the outflow volume remains relatively unchanged ($10\,639\text{m}^3$ and $10\,667\text{m}^3$ respectively) while the saturation excess value ($489\,096\text{m}^3$ and $541\,721\text{m}^3$ respectively) increases by an amount ($52\,625\text{m}^3$) which is equivalent to the error shown in Table 7-5 for this store. The imbalance for the overland store is absorbed into the soil store of the next cell through overland outflow, and this imbalance is transferred in turn to saturation excess for the overland store of this next cell. The storages that remain at the end of the time steps is also increased for the overland store, in comparing Table 7-5 and 7-6, due to the increases in input (saturation excess) to this store. The imbalances for the channel stores were relatively slight and these are absorbed into the channel outflows through the adjustment as well as increasing the remaining storage left over at the end. Overall the channel input and outflow is increased due to an increase in the overland contribution to this store. Besides increasing the remaining storages of the overland and channel stores, the majority of the imbalances are transferred by the adjustment to the cells downstream to be eventually felt at the catchment outlet as increased flow.

Through this test application, the operations of the model were verified to behave as expected. The continuity exercise allowed the opportunity to test possible solutions to the issue of imbalance. Although a solution was found through adjustment factors, it is felt that the resolution of this issue would be easier accomplished in the actual implementation of the model when evapotranspiration estimates could be used to absorb the imbalances.

(b) Hydrographs

Selected hydrographs are shown below to visually demonstrate how the model operates. The hydrograph for the soil, overland and channel stores of Cell 1 is shown in Fig. 7-24 (a, b and c), for the stores of Cell 2 in Fig. 7-25, for the stores of Cell 3 in Fig. 7-26 and for the stores of Cell 4 in Fig. 7-27. The corresponding computations of these hydrographs are given in Appendix D where the operations of the test model are shown. With due regard to the conventional representation of precipitation and streamflow hydrographs, the figures below show the inputs in each time step as pulsed

data (i.e. the input remains constant during the time step) and the outflow as break-point data (i.e. the measured flow at that instance at the end of the time step). In hydrology, the two principle variables of interest, namely precipitation and streamflow, are traditionally measured in this manner (Chow et al., 1988: 27). However, it must be remembered that in computing the intra- and inter-cell flows in the model setup, input into a store in each time step remained constant over the interval (as shown in the figures below) while a particular store's outflow was taken as the average of the outflows at the beginning of the time step and at the end of the time step. In this manner, this outflow could then become the constant input (over that time step) required for the receiving store or cell downstream.

In Fig. 7-24a (soil store of Cell 1), the chosen precipitation intensities over the 75 time steps can clearly be seen since the input for this store is purely from rainfall (since it is a source cell). The behaviour of the soil store as the regulating store of each cell can also be seen, i.e. the outflow from the store reaches a maximum upon the saturation of the store and the rest of the outflow forms the saturation excess. It is also evident from Figs. 7-24 to 7-27 that the soil store has the most retention capacity of all the stores (by considering the integral of the difference between inflow and outflow) and that overland and channel stores have little effect in the attenuation of the input.

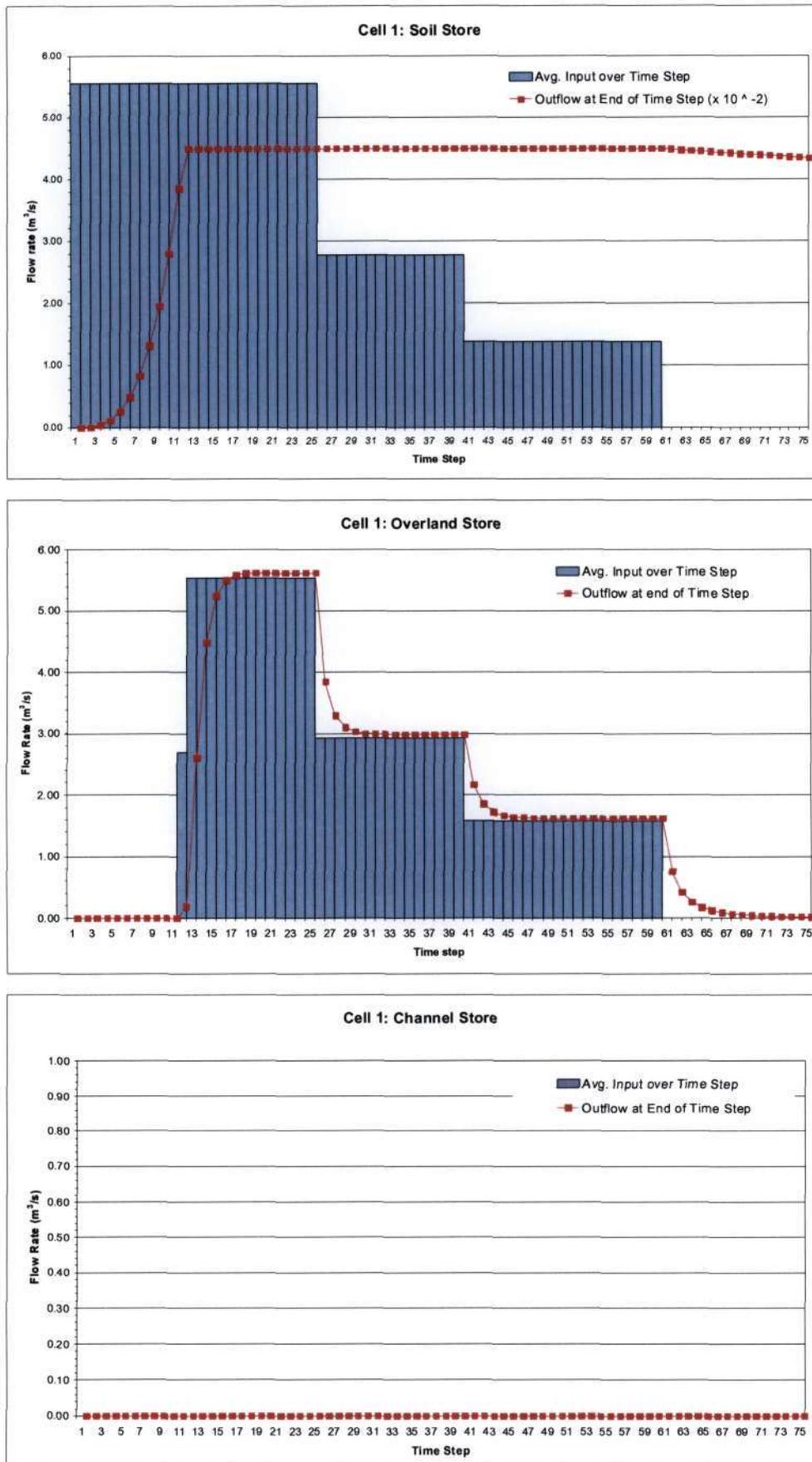


Figure 7-24a, b and c. The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 1.

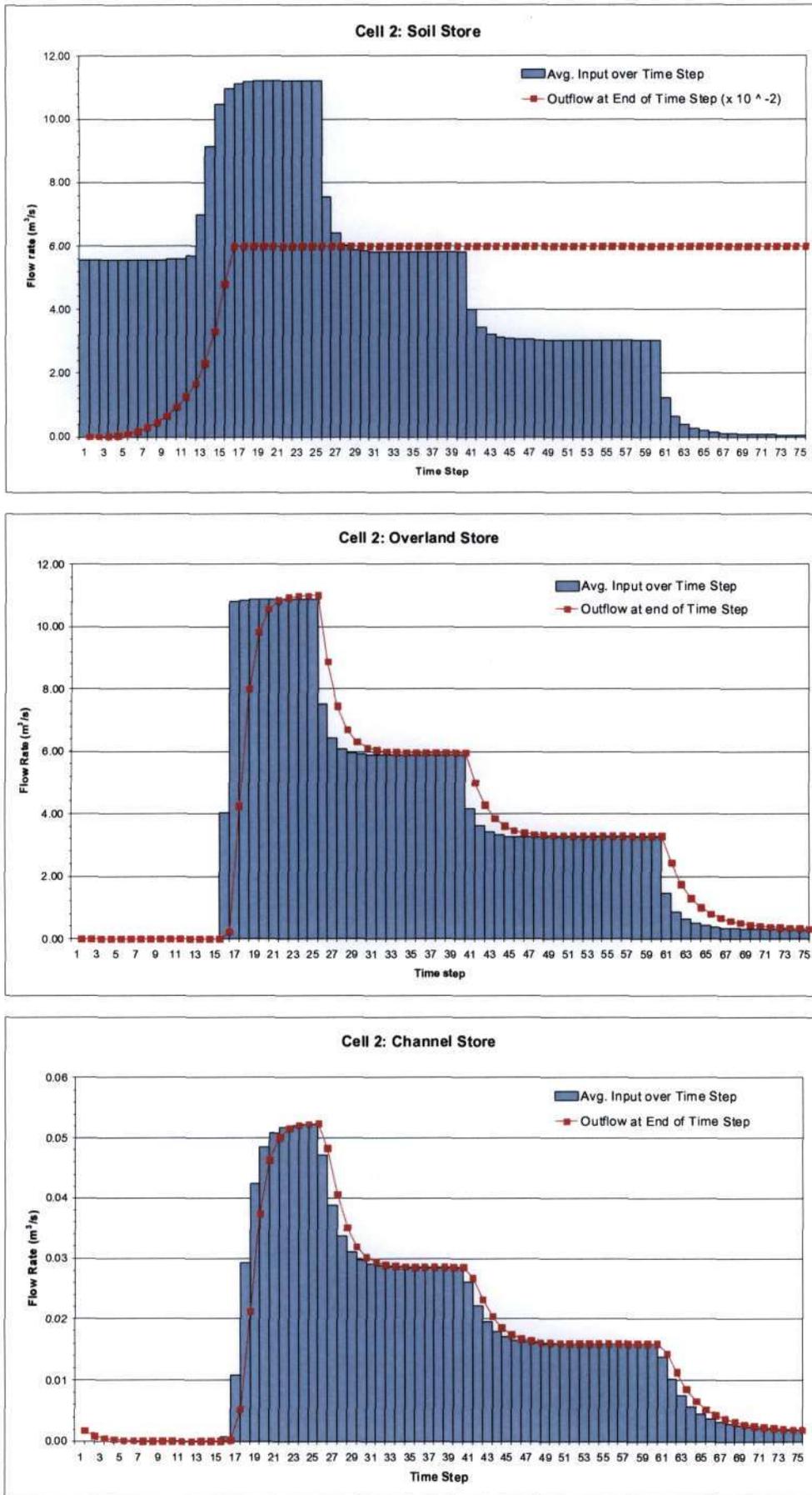


Figure 7-25a, b and c. The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 2.

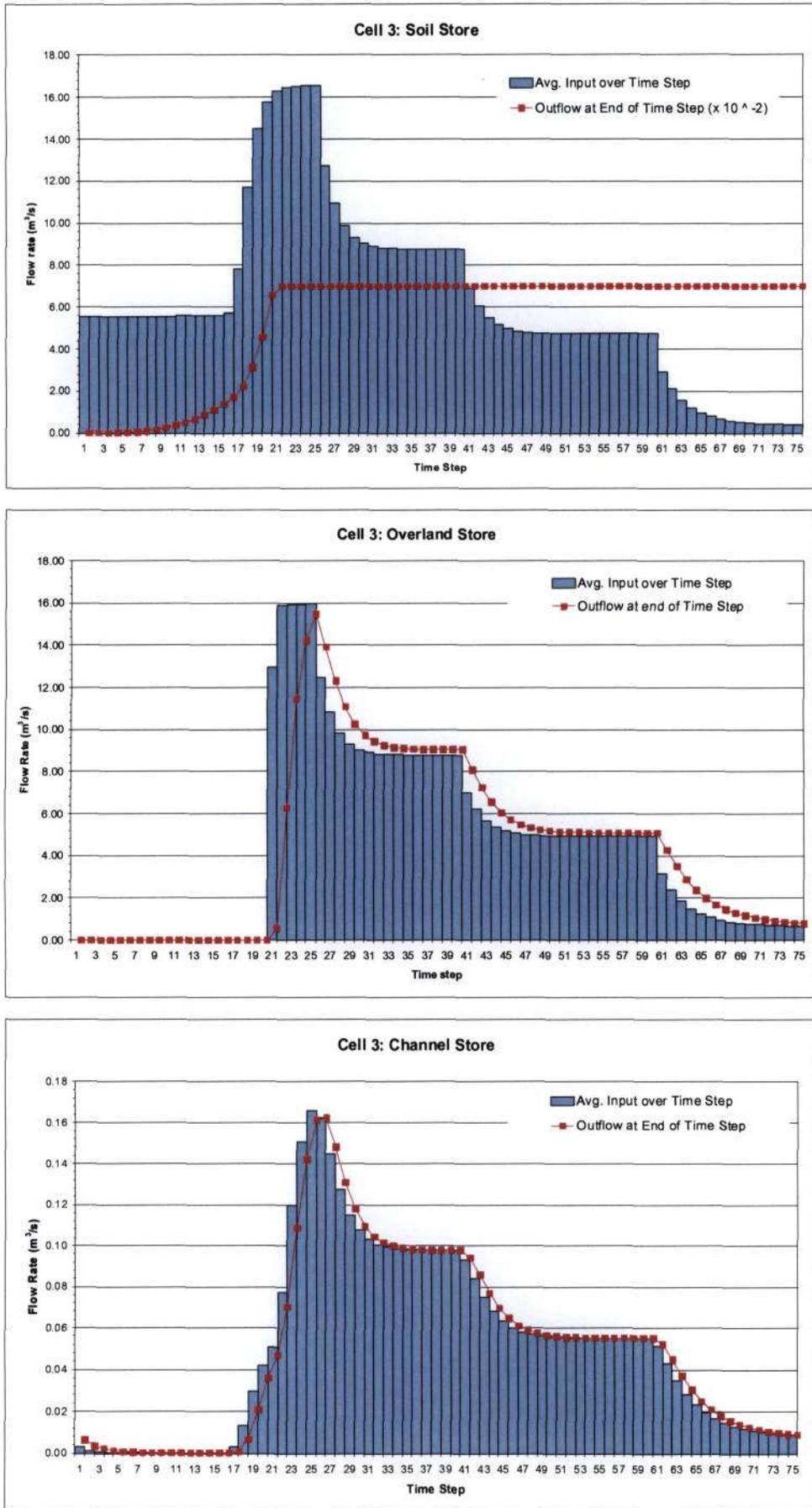


Figure 7-26a, b and c. The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 3.

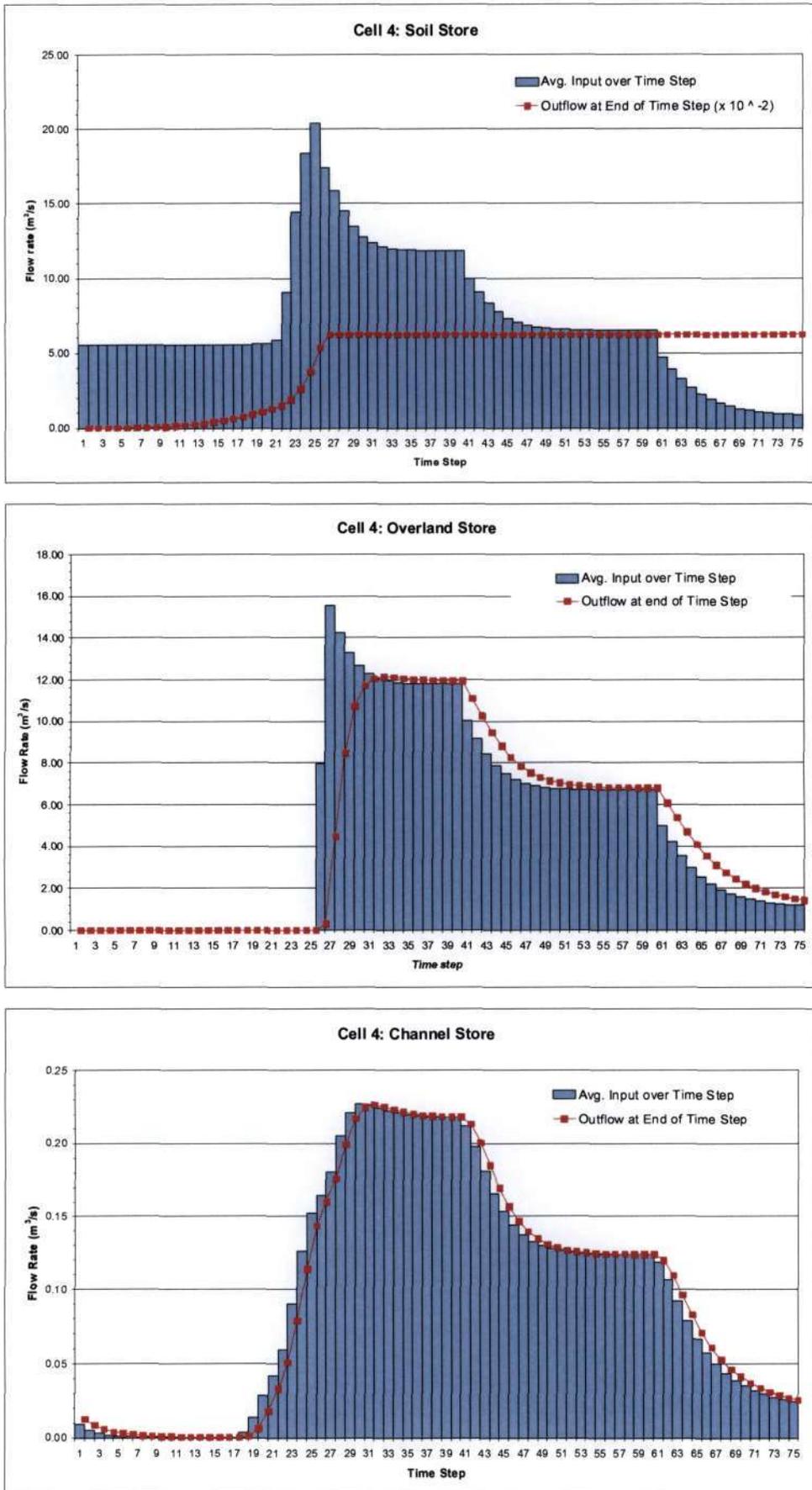


Figure 7-27a, b and c. The soil, overland and channel store's inflow and outflow hydrographs respectively for Cell 4.

7.3. Chapter summary

This chapter applied the TOPKAPI model, albeit in generic circumstances, to establish if the operations were correctly understood and if the model could be correctly implemented. This consisted of firstly gathering the required input data. These data then had to be managed and manipulated to suit the requirements of the TOPKAPI model. The Liebenbergsvlei catchment was chosen for this purpose since it is a highly instrumented catchment in South Africa. Thereafter, without actually using these data (but after having been satisfied that the data requirements of the TOPKAPI model were met), it was decided to test the operations of the model in a generic environment. With regard to this, a "four cell generic catchment" was created on a spreadsheet using Microsoft Excel™. This exercise allowed us to input rainfall into this "catchment", run the functions of the model and observe the outflow. This was done to verify if the model was operating correctly. A simple check of "continuity" was used to establish this. The model behaved as expected, and the issue identified in Chapter 6 and discovered in this chapter (namely the issue of a continuity imbalance) was rectified through simple adjustment factors.

It is envisaged that the exercises performed in this chapter would lay the groundwork and form the basis for the actual application of the TOPKAPI model for the Liebenbergsvlei catchment using a high level programming language such as C++. This latter aspect is left for a follow-up study. Other items which would need attention prior to the application of the TOPKAPI model, as a fully functioning hydrologic rainfall-runoff model, (and is not covered here) is the estimation of distributed evapotranspiration at the desired resolution for modelling the Libenbergsvlei catchment (1km square in this instance). Furthermore, the actual calibration of the static input data (terrain related) on the historic precipitation and streamflow records is also left for completion in the actual application of the model.

CHAPTER 8

8. CONCLUSION

Accurate, consistent and reliable flood estimates are needed by engineers responsible for the design of infrastructure that are at risk to flowing water and catchment / disaster managers that are responsible for the mitigation of flood damage. With regard to the former scenario, two design flood prediction methods were investigated in this research, namely the regional maximum flood (RMF) method in Chapter 3 and the rational formula method in Chapter 4. The outcome of these are summarised in Section 8.1.1 and 8.1.2 respectively.

This dissertation also focussed on the implementation of a fully distributed physically-based rainfall-runoff model for real-time flood forecasting applications. The TOPKAPI model was chosen for this purpose. Since this model has not been applied in South Africa, various sources of literature on the model were dissected, understood and explained in Chapter 6. Chapter 7 contains the results of the application the model, albeit in generic circumstances, to check the model's input requirements and operations.

This chapter summarises and discusses the contributions made in this research in Section 8.1 and concludes in Section 8.2 with recommendations for future studies.

8.1. Summary and discussion of research

8.1.1. A review of the regional maximum flood (RMF)

In Chapter 3 it was shown that the approximate return period of the regional maximum flood (RMF) estimate is 200 years. This was determined by simply comparing the RMF-based estimate with probabilistically modelled annual flood peak records. It was shown that the curves of the RMF envelope and the trend-line fitted to the 200-year modelled flood flows were approximately the same. The result of this investigation indicates that for all practical design purposes, it can be assumed that the RMF is equivalent to a 200-year flood.

It was also shown in Chapter 3 that the use of *catchment area*, as the sole landscape parameter in empirical equations (such as the RMF) provided the best predictions of floods when compared to empirical equations that included other landscape measures. The inclusion of other landscape parameters in addition to *area* did not improve the predictive ability of these methods. This was confirmed by using the coefficient of determination (R^2) to examine if the empirical model, with the inclusion of further landscape parameters, improved the predictive ability of the simple area-based model. Thus the use of *catchment area* as the sole independent variable in calibrating empirical equations seems practical (as it is easy to quantify) and efficient for the purposes of flood estimation.

8.1.2. The rational formula from the runhydrograph

In Chapter 4 the calibration of the rational formula's runoff coefficient was reviewed in order to assist in this parameters estimation and improve the reliability of this design flood prediction method. Characteristic design flood peak and volume pairs for certain catchments in South Africa were used for this calibration, offered by the runhydrograph method of Hiemstra and Francis (1979). The results produced calibrated coefficients that were of the same order of magnitude, but generally lower, than catalogued design values of this parameter from Chow et al. (1988: 498). It was found that the calibrated coefficients did not display any relationship with catchment characteristics. Thus, in validation, the coefficient values of Chow et al. were used as an upper bound estimate of the calibrated coefficients. The validation exercise showed a fair correspondence between the rational formula flood peaks (using the substitute coefficients from Chow et al.) and probabilistically modelled flood peaks for the validation catchments. This exercise confirmed the use of values from Chow et al. as upper bound estimates, since the rational formula flood peaks were larger than the probabilistically modelled flood peaks by a factor of 1.5 (on average). It was also discovered that the time base-length of the derived triangular hydrograph of this study was approximately between 1.9 and 2.6 times the catchment's time of concentration, depending on recurrence interval.

The results of this investigation were reasonably encouraging, in that the calibrated coefficients were scattered around published design values of the *c-coefficient* from Chow et al. (1988: 498). Although a proper validation of the calibrated coefficients was not done, in a sense a validation of the values of Chow et al. confirmed that these values are reasonable (although slightly conservative) for use in a design check for

large catchments as well as small. The hydrograph time base-length result was not conclusive as the observed variance was high.

8.1.3. TOPKAPI model

In Chapter 6, the TOPKAPI model was introduced and explained in great detail. This involved assimilating the literature on the model and related topics, dissecting it and disseminating it. The reasoning behind this was to prepare the ground for the eventual application of the model in a South African catchment such as the Liebenbergsvlei. In order to achieve this, the model's operations were tested, issues with regard to its application were identified and resolved, and the input data required by the model were gathered and prepared. The test application and resolution of issues was covered in Chapter 7 as well as the data preparation.

In Chapter 7, the data requirements of the TOPKAPI model were firstly identified. Thereafter the chapter proceeded with an explanation of how these data were gathered and processed for input into the model at the desired modelling scale, i.e. 1km². This involved predominantly work on a geographic information system (GIS) in order to gather and produce the information required. A digital elevation model (DEM) was used to calculate the surface slopes, areas drained, identify the flow pathways and detect the drainage networks of each cell/pixel. Other input requirements such as the soil data, landuse data and precipitation data was also gathered for each cell through a GIS and all the rasters were aligned with respect to the DEM. Furthermore, a special code was written to automatically trace the flow pathways as required by the TOPKAPI model, i.e. in only one of four possible directions from a cell. This was done because the GIS platform (ARCGIS™) used to perform the other operations described above did not operate in this manner and assigned flow from a cell in one of eight possible directions (a D8 model). The problems encountered with limiting a D8 flow direction raster to only four possible pathways (such as the formation of artificial sinks) were identified, explained and resolved with further code. It was also identified in Chapter 6 that readily available distributed evapotranspiration estimates for South Africa were lacking in terms of the input required for the model. A method was proposed on acquiring this information for the model, which is left for a follow up study. However, these estimates would also need to be included in the data sets gathered here for input into the TOPKAPI model.

Chapter 7 also described the test application of the TOPKAPI model. This entailed establishing a generic “four cell catchment” on a Microsoft Excel™ spreadsheet and running the model for 75 hourly time steps to verify that the operations of the model were correctly understood and implemented. A simple check to confirm this was to examine if *continuity* was maintained in the system. It was discovered that, due to the manner in which the analytical solution was derived and implemented, the solution had the tendency to underestimate the volume at the end of the time step on the rising limb of the store's hydrograph and overestimate this on the falling limb of the store's hydrograph. In testing the model, this problem manifested itself in the continuity checks, where imbalances were observed between the inflow, outflow, remaining volume and initial volume. In order to easily resolve this problem in this test setup, an adjustment factor was used to balance continuity. Although this had the desired result, it is felt that the resolution of this issue would be easier accomplished in the actual implementation of the model where evapotranspiration estimates could be used to absorb the imbalances. Despite this issue (of continuity), the test application confirmed the operations of the model to be valid.

The final step in achieving a fully functional rainfall-runoff model for the Liebenbergsvlei catchment would be to code the operations of the model using a high level programming language such as C++ based on the test application shown here and using the data gathered here. This is also left for a follow up study.

8.2. Chapter Summary

The research presented in this dissertation aimed to add some benefit to the engineering and hydrologic community with regard to flood estimation practices in South Africa for roads, bridges and dams. In this regard, two reviews of established design flood prediction methods were undertaken to assist in their use. The results are summarised in Sections 8.1.1 and 8.1.2. In addition to this, a pilot study was undertaken in introducing the TOPKAPI model for flood-forecasting purposes in South Africa. The contents of this study involved laying the groundwork for the models implementation in the Liebenbergsvlei catchment and is summarised in Section 8.1.3. Arising from this latter study, three issues will need to be addressed in follow-up studies prior to the models actual implementation. These are with regard to the acquisition of distributed evapotranspiration estimates as input for the model, the coding of the model using a high level programming language such as C++ and the

calibration of the model on past precipitation and flood flows for the Lienbenbergsvlei catchment.

REFERENCES

- Alexander W J R (1990) "Flood Hydrology for Southern Africa". SANCOLD, Pretoria.
- Alexander W J R (2002) "The Standard Design Flood – Theory and Practice". Report, Dept of Civil Engineering, University of Pretoria, Pretoria.
- Band L E (1986) "Topographic partition of watersheds with digital elevation models". *Water Resources Research*. Vol. 22, p. 15-24.
- Barnes H H (1967) "Roughness characteristics of natural channels". US Government Printing Office, Washington.
- Bartholmes J and Todini E (2003) "A combination of meteorological and hydrological models for flood forecasting applied to the upper river Po-basin". ACTIF 1st Workshop, Bologna.
- Basson M S, Allen R B, Pegram G G S, and van Rooyen J A (1994) "Probabilistic Management of Water resource and Hydropower Systems". Water Resources Publications. Colorado.
- Brooks R H and Corey A T (1964) "Hydraulic properties of porous media". *Hydrology papers*, No. 3. Colorado State University, Colorado.
- Cash J R and Karp A H (1990) "A variable order Runge-Kutta method for initial value problems with rapidly varying right hand sides". *ACM Transactions on Mathematical Software*. Vol. 16 (3), p. 201-222.
- Chilufya M S (2005) "Introductory GIS – The ARCGIS Way". Lecture Notes. School of Civil Engineering, Surveying and Construction. University of KwaZulu-Natal. Durban.
- Chow V T (1959) "Open-Channel Hydraulics – International Student Edition". McGraw-Hill, Tokyo.
- Chow V T, Maidment D R and Mays L R (1988) "Applied Hydrology". McGraw-Hill, New York.
- Crawford N H and Linsley R K (1966) "Digital Simulation in hydrology: Stanford Watershed Model IV". *Technical Report No. 9*. Department of Civil Engineering, University of Stanford. California.

-
- Cunnane C (1978) "Unbiased Plotting Positions--A Review". *Journal of Hydrology*. Vol. 37, p. 205-222.
- Dingman S L (2002) "Physical Hydrology". 2nd ed. Prentice Hall. New Jersey.
- De Michele C, and G Salvadori (2002) "On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition". *Journal of Hydrology*. Vol. 262 (1-4), p. 245-258.
- Dodson R D (1992) "Chapter 23: Advances in Hydrologic computation". Handbook of Hydrology. D R Maidment (ed.). McGraw-Hill, New York.
- Dunne T (1978) "Field studies of hillslope flow process". Hillslope Hydrology. M J Kirkby (ed.). p. 227-293. Wiley, New York.
- Doorenbos J and Pruitt W O (1977) "Crop water requirements". Irrigation and Drainage Paper 24. U. N. Food and Agriculture Organization, Rome.
- DLSI (1996) "A digital elevation model for South Africa". Directorate of Land Surveys and Information. Pretoria.
- Engman E T (1992) "Chapter 24: Remote Sensing". In: Handbook of Hydrology. D R Maidment (ed.). McGraw-Hill, New York.
- Everson C S (1999) "Evaporation from the Orange River: Quantifying open water resources". Water Research Commission. WRC Report No. 683/1/99. Pretoria.
- Francis D M (1979) "The Runhydrograph Applied". Master of Science Thesis. Civil Engineering, University of Natal, Durban.
- Francou J and Rodier J A (1967) "Essai de classification des crues maximales". Proceedings of the Leningrad Symposium on Floods and their computation. UNESCO.
- Freeze R A and Harlan R L (1969) "Blueprint for a physically-based, digitally-simulated hydrologic response model". *Journal of Hydrology*. Vol. 9, p. 237-258.
- GLCC (1997) "Global Land Cover Characteristics". United States Geological Survey (USGS). <http://edcdaac.usgs.gov/glcc>. Accessed 18 February 2005.
- Görgens A H M (2002) "Design Flood Hydrology". In: Design and Rehabilitation of Dams, G Basson (ed.). Institute for Water and Environmental Engineering, Department of Civil Engineering, University of Stellenbosch, South Africa. p. 460 - 524.
-

-
- Gradshteyn I S and Ryzhik I M (1980) "Table of integrals, series, and products". A. Jeffrey (ed.). Academic Press. Orlando.
- Gray D M (1973) "Handbook on the principles of hydrology". Water Information Centre. New York.
- Green W H and Ampt G A (1911) "Studies on soil physics, part I, the flow of air and water through soils". *Journal of Agricultural Science*. Vol. 4 (1), p. 1-24.
- Gumbel E J (1941) "The return period of flood flows". *The Annals of Mathematical Statistics*. Vol. 12 (2), p. 163-190.
- Haan C T (1985) "Lecture Notes". School of Bioresources Engineering and Environmental Hydrology. University of KwaZulu-Natal. Pietermaritzburg.
- Henderson F M (1966) "Open Channel Flow". Ch. 9, Flood Routing. The Macmillan Company, New York.
- Hiemstra L A V (1972) "Runhydrographs – A new technique in hydrograph generation". *Proc. International Symposium on Modelling Techniques in Water Resources Systems*. Ottawa, Canada, Vol. 1, p. 205-214.
- Hiemstra L A V (1973) "Runhydrographs for the sizing of dam spillways and the minimum reservoir capacities". *Proc. International Commission on Large Dams*. Madrid, Spain, Vol. 1, p. 217-237.
- Hiemstra L A V (1974) "Runhydrographs from Poisson-generated runlengths". *Journal of the Hydraulics Division, ASCE*. Vol. 100, p. 1617-1630.
- Hiemstra L A V and Francis D M (1979) "The Runhydrograph – Theory and Application for Flood Predictions". Water Research Commission, Pretoria.
- Hiemstra L A V, Zucchini W S and Pegram G G S (1976) "A method of finding the family of runhydrographs for given return periods". *Journal of Hydrology*. Vol. 30, p. 95-103.
- Horton R E (1932) "Drainage basin characteristics". *Trans. Am. Geophys. Union*. Vol. 13, p 350-361.
- Horton R E (1940) "An approach toward a physical interpretation of infiltration capacity". *Soil Science Society American Journal*. Vol. 5, p. 399-417.
- Horton R E (1945) "Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology". *Bull. Geol. Soc. Am*. Vol. 56, p 275-370.
-

- HRU Report No. 1/72 (1972) "Design Flood Determination in South Africa". D C Midgley (ed.). Hydrological Research Unit, Department of Civil Engineering, University of Witwatersrand, South Africa.
- <<http://www.cet.nau.edu/Projects/SWRA/research.html>> "Hydrologic cycle" (Accessed 10 May 2005).
- <<http://metsys.weathersa.co.za/simar-archive.html>> "Merged radar image 22/06/2005" (Accessed 25 June 2005).
- HYDRO1k (1996) "Elevation Derivative Database". United States Geological Survey (USGS). <http://edcdaac.usgs.gov/gtopo30/hydro>. Accessed 18 February 2005.
- Institute of Engineers Australia (1987) "Australian Rainfall and Runoff, A guide to Flood Estimation". D H Pilgrim (ed.). Canberra, Australia.
- Jenkinson A F (1955) "The frequency distribution of the annual maximum (or minimum) values of meteorological elements". *Quarterly Journal of the Royal Meteorological Society*. Vol. 81, p. 158-171.
- Jenson M E, Burman R D and Allen R G (1990) "Evapotranspiration and Irrigation Water Requirements". ASCE Manuals and Reports on Engineering Practice No. 70. American Society of Civil Engineers (ASCE). New York.
- Kirpich Z P (1940) "Time of concentration of small agricultural watersheds". *Civ. Eng.* Vol. 10 (6), p. 362.
- Kjeldsen T R, Smithers J C, and Schulze R E (2002) "Regional flood frequency analysis in the KwaZulu-Natal Province, South Africa, using the index-flood method". *Journal of Hydrology*. Vol. 255 (1-4), p. 194-211.
- Kovačs Z (1988) "Regional maximum flood peaks in Southern Africa". Technical Report: TR 137. Department of Water Affairs. Pretoria.
- Lettenmaier D P and Wood E F (1992) "Chapter 26: Hydrologic forecasting". Handbook of Hydrology. D R Maidment (ed.). McGraw-Hill, New York.
- Liu Z and Todini E (2002) "Towards a comprehensive physically-based rainfall-runoff model". *Hydrology and Earth System Sciences*. Vol. 6 (5), p. 859-881.
- Liu Z, Mario L V and Todini E (2005) "Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian catchment". *Hydrology and Earth System Sciences*. Vol. 9, p. 347-364.

- Mark D M (1988) "Network models in geomorphology". Chapter 4, Modelling Geomorphological Systems. M.G.Anderson (ed.). John Wiley & Sons, Ltd. p. 73-97.
- Marshall J S and Palmer W M (1948) "The distribution of raindrops with size". *Journal of Meteorology*. Vol. 5, p. 165-166.
- Martina M L V (2004) "The distributed physically based modelling of the rainfall-runoff process". PhD Thesis. University of Bologna.
- Martina M L V, Todini E and O'Connell P (2005) "Model lumping by incorporating hydrological processes derived from physically based models". EGU General Assembly presentation, 24 –29 April 2005, Vienna.
- Mazzetti C and Todini E (2002) "Combining raingauges and radar precipitation measurements using a Bayesian approach". In GeoENV IV – Geostatistics for Environmental Applications. J Gòmez-Hernàndez, A Soares and R Froidevaux (eds.). Kluwer Academic Publishers. The Netherlands.
- Midgley D C, Pitman W V and Middleton B J (1994) "Surface water resources of South Africa 1990 (WR90)". Water Research Commission. WRC Report No. 298/1/94. Pretoria.
- Monteith J L (1965) "Evaporation and Environment". Proceedings of the 19th Symposium of the Society for Experimental Biology. Cambridge University Press. New York. p. 205-233.
- Morel-Seytoux H J (1998) "Soil-aquifer-stream interactions – a reductionist attempt toward physical stochastic integration". *Journal of Hydrology*. Vol. 102, p. 355-379.
- Mulvaney T J (1850) "On the use of self-registering rain and flood gauges in making observations of the relations of rainfall and of flood discharges in a given catchment". *Proc. Inst. Civil Engineering Ireland*. Vol. 4, p. 18-31.
- MUSIC Final Report (2004) "Multi-sensor precipitation measurements integration, calibration and flood forecasting". Section 6, Deliverable no. 1.8, Feb 1st 2001 to July 31st 2004. E Todini (co-ordinator).
- O'Calaghan J F and Mark D M (1984) "The extraction of drainage networks from digital elevation data". *Computer Vision, Graphics and Image Processing*. Vol. 28, p. 328-344.

- Parak M (2003) "The Flood-Landscape Relationship". Unpublished undergraduate dissertation. School of Civil Engineering, University of KwaZuluNatal, Durban.
- Parak M and Pegram G G S (2006) "The rational formula from the runhydrograph". *WaterSA*. Vol. 32 (2), p. 163-180.
- Patton P C (1988) "Drainage basin morphometry and floods". *Flood Geomorphology*, Baker V R, Kochel R C and Patton P C (eds). John Wiley & Sons, Canada.
- Pegram G G S (2003a) "Rainfall, rational formula and regional maximum flood – some scaling links". Keynote Paper. *Australian Journal of Water Resources*. Vol. 7 (1), p. 29-39.
- Pegram G G S (2003b) "A national flood nowcasting system: Towards an integrated mitigation strategy". 11th South African National Hydrology Symposium. SANCIAHS. Port Elizabeth.
- Pegram G G S (2004) "Spatial Interpolation and Mapping of Rainfall (SIMAR), Volume 3: Data Merging for rainfall map production". Water Research Commission. WRC Report No. 1153/1/04. Pretoria.
- Pegram G G S and Sinclair D S (2002) "A linear catchment model for real time flood forecasting". Water Research Commission. WRC Report No. 1005/1/02. Pretoria.
- Pegram G G S and Parak M (2004) "A review of the regional maximum flood and rational formula using geomorphological information and observed floods". *Water SA*. Vol. 30 (3), p. 377-392.
- Pegram G G S, Sinclair D S and Wesson S (2005) "Daily rainfall mapping over South Africa: Mapping". Water Research Commission. WRC Report No. 1425 (in press). Pretoria.
- Penman H L (1948) "Natural evaporation from open water, bare soil and grass". *Royal Society of London Proceedings. Series A*, Vol. 193, p. 120-146.
- Petras V and du Plessis P H (1987) "Catalogue of Hydrological Parameters". Flood Studies. Technical Note No. 6, Department of Water Affairs, Pretoria.
- Pilgrim D H and Cordery I (1992) "Chapter 9: Flood Runoff". *Handbook of Hydrology*. D R Maidment (ed.). McGraw-Hill, New York.
- Philip J R (1957) "The theory of infiltration: 1". The infiltration equation and its solution. *Soil Science*. Vol. 83, p. 345-357.

- Reed D W (1984) "A review of British flood forecasting practice". Institute of Hydrology. Report No. 90.
- Rooseboom A, Basson M S, Loots C H, Wiggot J H and Bosman J (1981) "National Transport Commission road drainage manual". 1st ed. National Transport Commission, Directorate Land Transport, Pretoria.
- Richard L A (1931) "Capillary conduction of liquids through porous mediums". *Physics*. Vol. 1, p. 318-333.
- SAICE (2004) "Code of ethics for members". The South African Institution of Civil Engineering.
- Schmidt E J and Schulze R E (1987a) "SCS-based design runoff". Water Research Commission. Pretoria.
- Schmidt E J and Schulze R E (1987b) "SCS-based design runoff – User Manual". Water Research Commission. Pretoria.
- Schmidt E J, Schulze R E and Dent M (1987) "SCS-based design runoff – Appendices". Water Research Commission. Pretoria.
- Schulze R E (1995) "Hydrology and Agrohydrology". Water Research Commission. WRC Report TT69/95. Pretoria.
- Soil Conservation Service (1972) "SCS Method". National Engineering Handbook, Sec. 4, Hydrology. U.S. Department of Agriculture. Washington.
- Singh V P and Woolhiser D A (2002) "Mathematical modeling of watershed hydrology". *Journal of Hydrologic Engineering*. Vol. 7 (4), p. 270-292.
- Smithers J C and Schulze R E (2001) "Design Runoff Estimation: A review with reference to practices in South Africa". Tenth South African National Hydrology Symposium. Pietermaritzburg.
- Smithers J C and Schulze R E (2003) "Design Rainfall and Flood Estimation in South Africa". Water Research Commission, WRC Report 1060/01/03.
- SIRI (1987) "Land Type Series". Department of Agriculture and Water Supply, Soil and Irrigation Research Institute. Memoirs on the Agricultural Natural Resources of South Africa. Pretoria.
- Strahler A N (1952) "Hypsometric (area-altitude) analysis of erosional topography". *Bulletin of the Geological Society of America*. Vol. 63, p 1117-1142.

- Strahler A N (1964) "Quantitative geomorphology of drainage basins and channel networks". V T Chow (ed.). Handbook of applied hydrology, Sec. 4-11. Mcgraw Hill, New York.
- Shreve R L (1966) "Statistical law of stream numbers". *Journal of Geology*. Vol. 74, p. 17-37.
- Tarboton D G, Bras R L and Rodriguez-Iturbe I (1991) "On the extraction of channel networks from digital elevation data". *Hydrological Processes*. Vol. 5, p. 81-100.
- Todini E (1995) "New trends in modelling soil processes from hillslope to GCM scales". The role of Water and the Hydrological Cycle in Global Change. H R Oliver and S A Oliver (eds.). Nato ASI Series: I, Global Environmental Change, 31, p.317-347.
- Todini E (1996) "The ARNO rainfall-runoff model". *Journal of Hydrology*. Vol. 175, p. 339-382.
- Todini E (2005) Personal communication with supervisor at European Geosciences Union (EGU) General Assembly, May 2005. Austria.
- Weibull W (1939) "A statistical theory of the strength of materials" Ingeniors Vetenskaps Akademien. Proceedings No. 51, p. 5-45.
- Wood E F, Sivapalan M, Beven K J and Band L (1998) "Effects of spatial variability and scale with implications to hydrologic modeling". *Journal of Hydrology*. Vol. 102, p. 29-47.

A review of the regional maximum flood and rational formula using geomorphological information and observed floods

Geoff Pegram* and Mohamed Parak

Civil Engineering Programme, University of KwaZulu-Natal, Durban 4041, South Africa

Abstract

Flood estimation methods in South Africa are based on three general approaches: empirical, deterministic and probabilistic. The "quick" methods often used as checks are the regional maximum flood (RMF) and the rational formula (RF), which form part of the empirical and deterministic methods respectively. A database of annual flood peaks was used in a probabilistic approach to review these methods and to provide preliminary insight into their estimates of flood peaks. This paper examines the following: the relationship between floods and landscape; the estimation of the return period of the RMF; the use of ratios in scaling RMF flood peak estimates to flow rates of shorter return periods; the applicability of the modified rational formula (MRF); the examination of the relationship between scaling parameters and regional parameters. It turns out that the RMF is the best of all methods examined in this preliminary study (other than statistical) in estimating the 200-year flood peak at an ungauged location.

Keywords: flood estimation, rational formula, regional maximum flood, generalised extreme value distribution

Introduction

The realistic estimation of the magnitude of a design flood peak with a chosen probability of exceedence that can be expected at a given site in a given region is fundamentally important in the planning, design and operation of hydraulic structures and for the preservation of human life and property. The basic approaches involved in flood estimation are the empirical, deterministic and probabilistic approaches. These methods are calibrated from historical flood records from gauged catchments and their relative usefulness depends on the accuracy with which they are able to predict flood sizes in ungauged catchments. In South Africa, reasonable estimates of maximum recorded flood magnitudes are derived from the use of the empirically-based approach of the regional maximum flood (Kovacs, 1988), and design floods may be determined using deterministic approaches such as the rational formula (RF), the SCS model or the unitgraph method and from the analyses of flood frequencies through a probabilistic approach.

Kovacs' **empirical method** is probably the most robust method available locally and, relatively accurately, predicts the regional "maximum" flood that can be expected from a given site based only on the site's catchment area and region. The advantage of the empirical method is its ease of use as it deals directly with the parameter of interest, namely the flood peak discharge, and avoids the assumptions involved in transforming rainfall inputs into flood outputs. The disadvantages of the RMF method are that:

- The recurrence interval (RI) associated with this "maximum" is not clear, although Kovacs estimated it to be greater than 200 years
- The regions defined by individual K -values have widely varying rainfall properties and
- It seems naive to estimate flood peaks on area and zone only.

The **deterministic** rational formula (RF) approach involves (in a simple, but sound manner) the analysis of all the factors involved in flood prediction from converting rainfall inputs into flood outputs; it usually carries a caveat that it should not be used for "large" catchments, but recent work (Alexander, 2002 and Pegram, 2003) has shown that this caution is too conservative.

Flood **frequency analysis** involves the fitting of a probability model to the sample of annual flood peaks, recorded over a period of observation, for a catchment of a given region. The model parameters established can then be used to predict the extreme events of large recurrence interval. The advantage of this method is that the events of large recurrence interval, which are longer than the record period, can be determined through cautious extrapolation of the fitted distribution based on the model parameters. The disadvantage of this method is that it can only be applied where data have been collected and it is often not clear how the analysis can be extended to ungauged locations.

The question that arises is "which method is fair to use?" The answer depends on the availability of data. When no hydrological (rainfall and runoff) records exist for a catchment, the empirical methods provide the only means of flood prediction. This situation is the most common case in the design of hydrological projects. When estimates of design rainfall are available (Adamson, 1981; Smithers and Schulze, 2003) or rainfall records suitable for a frequency analysis are available from a nearby rain-gauge, then the rational formula (RF) becomes applicable, in addition to the empirical. When flood records of sufficient length (>30 years or so) exist, possible future flood peaks of given frequency can be determined by modelling past floods with an extreme value distribution. Even in this fortunate situation, it is prudent to crosscheck the frequency estimate with deterministic and empirical estimates.

It is the aim of this exploratory study to provide a review of the above methods in order to determine the accuracy of the estimates of the design flood, where the design flood is the flood associated with a chosen return period or recurrence interval of exceedence. The base data are the set of annual flood peak records from 130 sites around South Africa that were used *inter alia* by Kovacs (1988) in his empirical study.

* To whom all correspondence should be addressed.
 ☎ +2731 260-3057; fax: +2731 260-1411; e-mail: peggram@ukzn.ac.za
 Received 8 January 2004; accepted in revised form 14 May 2004.

To summarise: this paper attempts to provide preliminary insight into the following questions concerning the RMF and RF flood determination methods in South Africa using the recorded flood peaks:

- Does the addition of landscape data (catchment morphometry) improve the prediction of floods by the RMF?
- Can a return period be associated with the RMF by comparing its computed magnitude with those modelled from historical records?
- Are simple country-wide Q_r/Q ratios valid for scaling flood maxima (or RMF values) to floods of shorter (or even longer) return periods?
- Is the modified rational formula (MRF) a useful modification and reasonable alternative to the RF and other flood prediction methods?
- Are there any inferences that can be drawn from the variation of the shape parameters k of the GEV Distribution, used to model the observed floods, and Kovacs' regional K -values?

The methodologies involved in assessing each of the objectives listed in this paper will be expanded in detail in the sequel. Before this can be done, an explanation of how the recorded data set was used in the calibration and validation of the objectives outlined is given.

The use of recorded flood data in this study

Annual flood peaks from 130 catchments across South Africa were obtained from Zoltan Kovacs of the Department of Water Affairs by Peter Adamson while working with the first author in 1988 and 1989. This data set, although old (final year of record was 1988), provided the starting point for this pilot study in the review of these flood determination methods. The length of record of the data set used herein ranged from 9 years to 76 years and forms a sub-set of the data used by Kovacs for the construction of the RMF curves. To find the return period associated with each annual peak, the Weibull Plotting Position was used (it is more conservative than the Cunnane Plotting Position), which is expressed as:

$$T = \frac{N+1}{r} \quad (1)$$

where:

- T is the return period (years) associated with the flood peak of rank r
- N is the length of record (years)
- r is the rank of the flood peak; $r = 1$ for the largest peak.

This resulted in a list of annual peaks each with an associated return period for each catchment. Following the work of De Michele and Salvadori (2002) and Kjeldsen et al. (2002), the distribution of these peaks was assumed to follow a generalised extreme value (GEV) distribution. This distribution takes the following form:

$$Q_T = \mu + \sigma y_T \quad (2)$$

where:

- Q_T is the T -year return period flood peak estimate
- μ, σ are shift and scaling parameters respectively
- y_T is the GEV reduced variate corresponding to a T -year return period, i.e.

$$y_T = \frac{1}{k} \left[1 - \left\{ -\ln \left(1 - \frac{1}{T} \right) \right\}^k \right] \quad (3)$$

where:

- k is a shape parameter. When $k = 0$, the GEV reduces to the EVI or Gumbel distribution.

This model of the flood data formed the basis with which to review the other approaches. Some of these data and their distribution fits are presented in Table A1 (Part 3) in the Appendix.

Empirical approach extended by including landscape properties

In his empirical approach Kovacs (1988) determined envelopes of the maximum flood peaks from the original extended data set, of which, as has been noted above, the data in the Appendix are a subset. Kovacs' data set included some rare singleton floods (not used in this study) to which he cautiously ascribed a representative record length not exceeding 200 years. He used this extended set to obtain the RMF lines based on the Francou-Rodier equation. The technique was to plot maximum flood peaks against catchment area for hydrologically homogeneous regions to produce envelope curves which define the upper limit of expected flood peaks for a given region. The curves are defined by the following equation:

$$Q_{RMF} = 10^6 \left(\frac{A}{10^8} \right)^{1-0.1K} \quad (4)$$

where:

- A is catchment area in km^2
- K is a regional dimensionless factor which accounts for the influence of variations in rainfall, geology, land-form and vegetation cover in flood production.

It should be noted at this juncture, that the "secret" to the success of the RMF is the careful way in which Kovacs chose the regions to group the flood data. He did this by examining the actual K -value (from Eq. (4)) for each catchment where the flood peaks and catchment areas were known. Regional boundaries of K were delimited by considerations of individual K -values within the region, the number and accuracy of the data in a particular area, existing boundaries, maximum recorded 3 day storm rainfall, topography, catchment orientation with respect to dominant storm generating weather systems, general soil permeability, main drainage network and the location of large dams situated upstream from the gauging sites (Kovacs, 1988). Of these considerations, individual K -values were evidently the most important and the regions were traced based on this. In areas of high flood peak potential a difference of 0.2 between individual K -values was allowed for and a difference of 0.6 in areas of low flood peak potential.

What is evident from Eq. (4), and all other derived empirical equations produced for the prediction of floods, is its dependence on *Area* as an independent variable. Because of the RMF's apparent naivety, one might expect other parameters of the fluvial landscape to play an important role in flood response and make the estimates more accurate. Flood geomorphologists, such as Horton (1932; 1945) and Strahler (1952; 1964) and many others since have been interested in relating flood discharges to physical measures of landscape (morphometry). They identified parameters of the fluvial landscape which intuitively would correlate well with flood discharge.

Linear parameters (such as stream orders and stream lengths), areal parameters (such as catchment area, catchment shape and drainage density) and relief parameters (such as catchment relief, catchment slope, channel slope and ruggedness number) are some of the physical measures that have been identified as significantly

affecting flood response. One can expect such a relationship between flood discharge and catchment morphometry to exist because a catchment is effectively "an open system trying to achieve a state of equilibrium" (Strahler, 1964). Precipitation is input to the system and soil (eroded material) and excess precipitation leave the system through the catchment outlet. Within this system an energy transformation takes place converting potential energy of elevation into kinetic energy where erosion and transportation processes result in the formation of topographic characteristics. Thus it is evident that floods, and the landscape through which they drain, form a mutual relationship and ultimately catchment morphometry should reflect this phenomenon. In this pilot study, an effort is made to determine if landscape parameters improve the prediction of floods in empirical equations based solely on catchment area.

What is the recurrence interval of the RMF?

What is also evident from the RMF method of flood determination is that one is not easily able to associate a return period with the estimated floods. The envelope floods (estimated from the RMF lines) have been described as the maximum flood that the site has experienced. This is not easy to quantify in terms of a return period. Kovacs himself estimates the return period to be greater than 200 years (Kovacs, 1988), although he does not explicitly model their probability distribution. Where the representative period (N) of a flood was not known, Kovacs did not allow this to exceed 200 years and a provisional N value was estimated based on the assumption that the ratio of the 200-year peak to RMF, Q_{200}/RMF was 0.65.

When determining a design flood, the exact magnitude of the flood and its probability of exceedance need to be known. The absence of an estimate of the return period associated with the RMF makes the quantification of risk by this method problematic and, as it represents maximum discharges, it tends to be used by designers as a conservative method. This article aims to, *inter alia*, determine a return period associated with the RMF by simultaneously plotting the floods determined from the RMF method and the historical floods extrapolated to the 50-, 100- and 200-year recurrence intervals modelled with the GEV distribution.

The first author has for many years suggested that the RMF envelopes have a recurrence interval of about 200 years, as estimated by the Weibull Plotting Position. This estimate was based on the following argument: the data used by Kovacs (1988) in the construction of the RMF lines had, in many instances, record lengths (actual and estimated) of the order of 100+ years. The RMF lines are envelopes, drawn above the data whose maximum record length N was 200 years. If we are conservative and estimate the recurrence interval of the RMF line using the Weibull Plotting Position, the RI (T) of the largest observation would be $T_N = (N+1) \approx 200$ years. It was decided to examine this conjecture as part of this study.

The use of Q_T/Q_2 ratios for scaling flood maxima

It is useful to know how to scale the "200-year RI" RMF or any other flood of recurrence interval T -years to shorter return period floods where desired. The first author suggested such a scaling in Chapter 2 of TRH 25 (1994). It was thought that this study was also an opportune time to check that assumption which was based on the following argument.

Hiemstra and Francis (1979) examined the relationship between the peak flood discharge of a catchment and its hydrograph

T	2	10	20	50	100	200	1000	10000
Q_T/Q_2	1	3.57	5.18	7.80	10.24	13.14	22.00	41.24

shape defined by the volume. What they discovered was that for extreme events, the peak discharges of various magnitudes were well modelled by the censored log-normal distribution. They extracted the statistics of many floods in the Department of Water Affairs and Forestry's break-point continuous flow rate database at that time and found the coefficient of variation of the peak discharges averaged 1.3 with a fairly small variation. Based on this, the first author produced ratios which relate the T -year flood to the 2-year flood. These ratios, Q_T/Q_2 , reproduced from TRH 25 (1994) in Table 1, enable one to convert any flood of a given RI to a T -year flood.

To check this assumption in this study, the maximum observed flood recorded in the observation period from each of the 130 catchments was associated with a return period using the Weibull Plotting Position ($T_{max} = N+1$). This flood was then scaled to 10- and 50-year flow rates using the Q_T/Q_2 ratios and compared with those computed from the GEV model fitted to the full set of data in each record. These values were then compared and it was determined if these ratios are applicable in reducing flood maxima to floods of desired recurrence intervals.

The modified rational formula (MRF)

The rational formula is expressed (in S_I units) as:

$$Q_{peak} = ciA/3.6 \quad (5)$$

where:

- c is a dimensionless runoff coefficient which ranges from 0 to 1
- i is the intensity of rainfall (mm per hour) of return period T (years) and duration T_c , where T_c is the time of concentration (hours) of the catchment
- A is the area of the catchment (in km^2).

This formula is usually limited to catchments with small areas ($< 100 km^2$). The reason usually given for this is that the formula does not take into account the areal reduction factor (ARF) and utilises point design rainfall intensity. It should be noted that flood-causing rainfall in smaller catchments is mainly due to concentrated thunderstorm activity, whereas flood-producing rainfall in larger catchments is mainly due to long-duration, widespread synoptic events (Pegram, 2003). The consequence is that the larger the catchment, the longer the duration of the flood-causing rainfall. To simplify the analysis, Pegram (2003) used the scaling properties of the GEV distribution fitted to rainfall depths, hence, using the GEV distribution defined in Eq. (3), the precipitation scaling relationship becomes:

$$P_{d,T} = (\mu + (\sigma/k) \{1 - \{-\ln(1-1/T)\}^k\}) \sigma^{1-\eta} \quad (6)$$

where:

$P_{d,T}$ is the rainfall depth of duration d and return period T .

For each of Kovacs' regions, representative 1-, 2- and 3-day rainfall depths for 2-, 5-, 10-, 20- and 50-year return periods were extracted

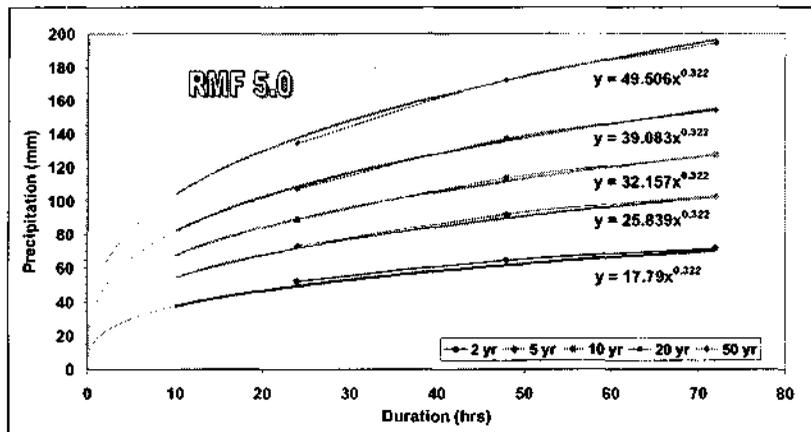


Figure 1
Fit based on the GEV (smooth curves), to an average of Adamson's (1981) data (dots) for Kovacs (1988) RMF Region 5 (from Pegram, 2003). The thin lines are trend-lines fitted to each set of the 1-, 2- and 3-day rainfall duration data. The thick lines are the combined models fitted to all the data with a common power law relationship.

from Adamson (1981) by Westray (2001). These were averaged (pooled) by region and Eq. (6) was fitted to the 15 points by Least Squares. An example is given in Fig. 1 where the pooled data and the fitted function are compared for Region 5. The k and η values were fixed at -0.182 and 0.678 respectively by using the whole South African data set as a first approximation (Pegram, 2003). Values of μ and σ were the parameters that were estimated for each region. It was found that the coefficient of variation $C_v = \sigma/\mu$ was effectively independent of Kovacs' regions, so the major variable to concentrate on was the parameter μ .

In addition to this simplification, for all the catchments whose data are available in the report by Petras and Du Plessis (1987), the time of concentration T_c computed from the Kirpich (1940) formula: $T_c = 0.0633(L/S)^{0.385}$ (where as usual, L is the length and S is the average slope of the catchment's longest watercourse) was set to the duration of the flood-causing storm as demanded by the RF. When this duration T_c was plotted against area, the points clustered around a curve to which a power law relationship could be fitted. This is also the practice in Australian Rainfall and Runoff (AR&R, 2001). For interest sake, this was superimposed on the areal reduction factor (ARF) diagram, published in the Flood Studies Report (FSR, 1975), which appears as Fig. 2. It is possible that the FSR's ARF curves over-estimate the ARF in South Africa, but the degree is likely to be a matter of climate. Conscious of this, it is still remarkable that the $Area \sim T_c$ curve yields an almost constant ARF value of 88% across the FSR curve. Thus, as long as the precipitation intensity used in the rational formula corresponds to the time of concentration of the catchment, the point rainfall is automatically scaled by a constant ARF. Combining these ideas, the MRF was then expressed (Pegram, 2003) in preliminary form as:

$$Q_{MRF} = c \times 0.318 P_{1day,2} [1 + 0.385 y_T] A^{0.558} \quad (7)$$

where:

- c is the conventional rational formula (RF) $c: 0 < c < 1$
- $P_{1day,2}$ is the median 1d annual maximum rainfall available from maps (e.g. Adamson, 1981; Smithers and Schulze, 2003)
- y_T is the reduced variate of the GEV Distribution of the rainfall
- A is the catchment area in km^2 .

In this paper the 10-, 20- and 50-year floods of the MRF are compared with the observed flood peaks modelled with the GEV distribution of the same recurrence intervals. The intention is to determine whether the MRF in its coarse form is possibly a useful candidate for predicting the design floods of a catchment.

Does the GEV regionalise following the RMF?

The annual observed flood data series, extracted from the observed records, were modelled using the GEV distribution. This was explained above. The records were thought to be long enough, in most cases, to make reasonable predictions of future events. Following this analysis, it was of interest to determine if the shape parameter k established by modelling historical floods using the GEV distribution, display any trends with a region descriptor such as Kovacs' regional K -value. That concludes the introduction. The full analyses are reported in the following sections.

Floods and landscape

Landscape data from 25 catchments were extracted in a preliminary study by Parak (2003) that corresponded with the peak discharges of the catchments modelled in this study. Parak (2003) captured morphometric data of 45 catchments across the country in his investigation into the relationship between floods and landscape. He used already catalogued data (Petras and Du Plessis, 1987 and Kovacs, 1988) and supplemented this with further data through map work from Midgley et al. (1994). In this paper the landscape data were utilised to assess whether they improved the prediction of floods compared with the RMF, which uses only catchment areas in particular regions. The flow rate that was used for comparison here was the 20-year event determined by modelling the historical floods of the catchments using the GEV distribution, the rationale being that:

- It would be the least likely estimate to be affected by fitting the wrong probability distribution
- Many of the records were longer than 20 years.

The flood and landscape data were split into two groups, one for calibration and the other for validation. The landscape data included catchment area, mean channel slope, mean annual precipitation, drainage density, catchment relief and ruggedness number. These are summarised in Table A1 and explained in the Appendix and a typical catchment and its derived geometry are shown in Fig. 3 (from Parak, 2003). It is acknowledged that the landscape data catalogued are sensitive to map scale, i.e. at different scales, different values of the parameters will be obtained. For example, the river detail shown on a larger scaled map is much less than that which is shown on fine-scaled maps. This has a direct influence on the magnitudes of the landscape parameters. Measures such as total stream length, stream orders, drainage densities and ruggedness

Figure 2
The FSR diagram for ARF (FSR, 1975), as contours in percentages, with T_c vs area relationships (using Kirpich's (1940) formula:
 $T_c = 0.0633 (L/S)^{0.385}$ for South African catchments superimposed (Westray, 2001). The best fit is:
 $T_c = 0.148 A^{0.651}$.

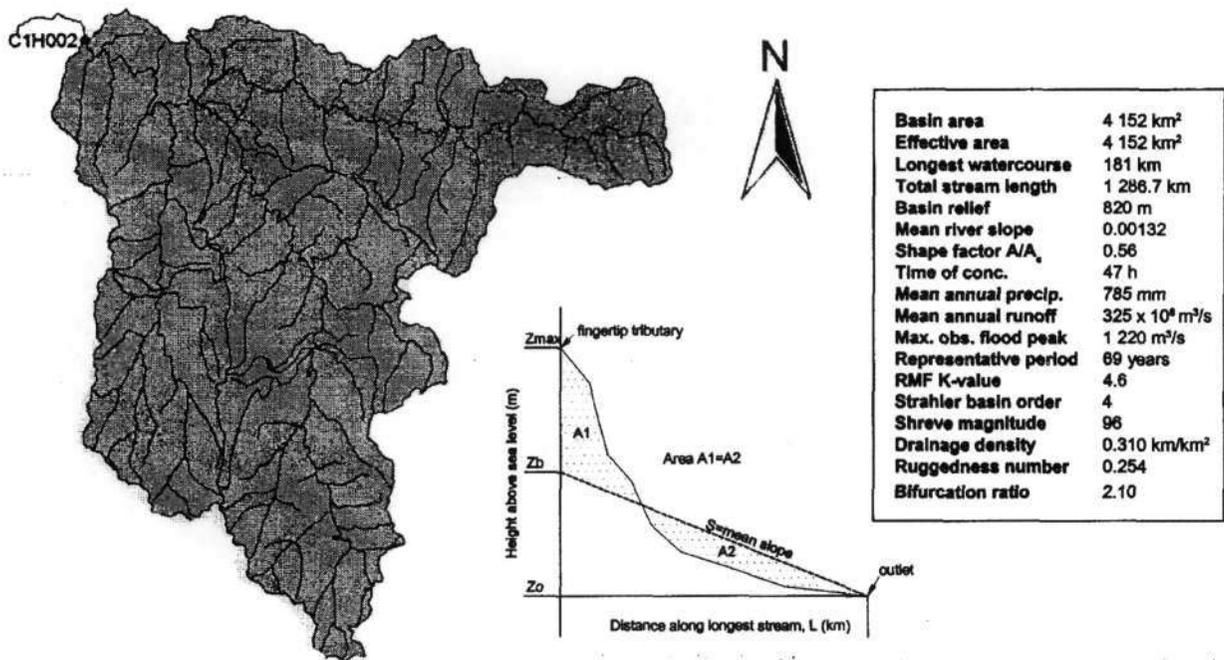
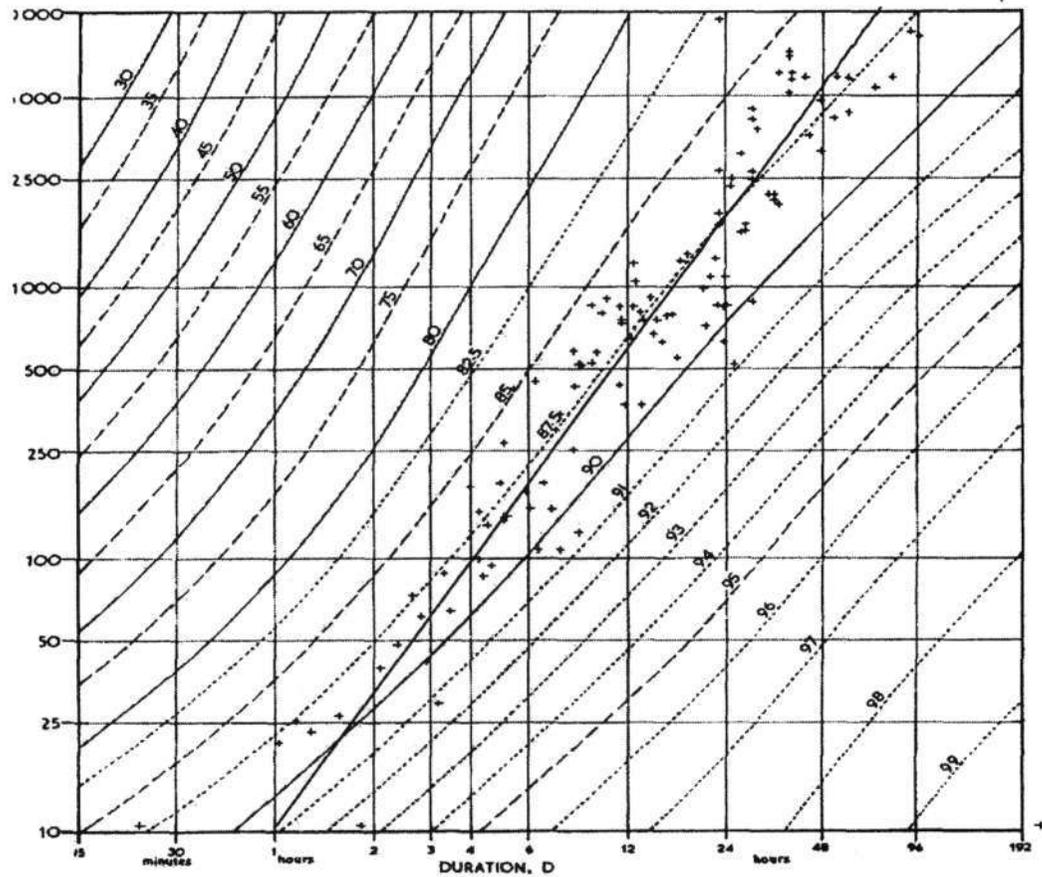


Figure 3
Plan, long section and basin properties of the Klip River catchment (represented by gauge C1H002) in the eastern highveld area of South Africa (Petras and Du Plessis, 1987; Kovacs, 1988; Midgley et al., 1990 and Parak, 2003). Reference should be made to the Appendix for the definitions of these parameters and those summarised in Table A1.

numbers are all dependent on the scale of the map from which these parameters were extracted. More accurate measures can be made with the use of finer scaled maps, but this comes at the expense of greater effort and time requirements. Parak (2003) used uniform scaled maps from Midgley et al. (1990) showing river detail at 1:250 000 for the data extraction.

The criterion for choosing an appropriate model was based on the determination of the R^2 statistics through stepwise regressions. The first group of flow rates were plotted against catchment area to determine a regression equation and R^2 statistic in the calibration set. The regression equation was then used to generate flow rates of the second (validation) group from the independent variable and these estimates were plotted against the recorded ones of the same group to see if they mimicked each other. The degree of validation was based on the strength of the R^2 statistic. Subsequently, other landscape data were combined with catchment area to examine if they improved the strength of the relationship (based on the R^2 statistic) in calibration and validation. A conclusion was drawn based on the examination of the R^2 statistic in calibration and validation of the two groups of flood and landscape data.

In the original study Parak (2003) examined the relationship between the flood peaks and the various candidate landscape parameters. The model, given by Eq. (8), was selected after examining the literature on geomorphological estimates of floods and carefully plotting pairs of variables. A power-law relationship was selected and various groupings of "independent" variables were included in the regression equation, which was the logarithm of Eq. (8), shown as Eq. (9). The model selection process was performed by fitting the model to a calibration set and checking the fit for a validation set. The most suitable formulation was a power relationship of the form:

$$Q_{20} = aA^b X^c Y^d \dots \quad (8)$$

where:

a, b, c and d are parameters to be regressed from the data and A, X and Y are landscape quantities.

The formulation for regression was to take logarithms of Eq. (8) and regress using the linearised model:

$$\log(Q_{20}) = \log(a) + b \cdot \log(A) + c \cdot \log(X) + d \cdot \log(Y) \quad (9)$$

Figure 4 shows the calibration of the empirical equation defining the 20-year flood as a function of catchment area. The R^2 statistic from this model implies a strong relationship (0.856) and good fit. When this empirical model is tested against the reserved data of the second group in validation (Fig. 5), the fit is evidently poor, producing a moderate R^2 statistic of 0.538. When further landscape data are added to catchment area in the hope of improving the fit of the empirical models, the results are no better. The addition of landscape data as independent variables in the prediction of floods did not improve flood prediction and it seemed as though the best model of floods and landscape is simply area based. These results

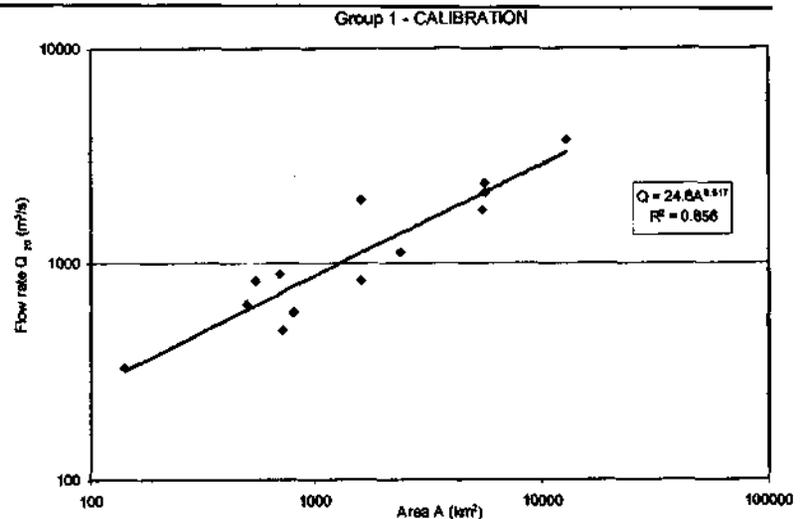


Figure 4
 Q_{20} vs. area - calibration set for 13 catchments

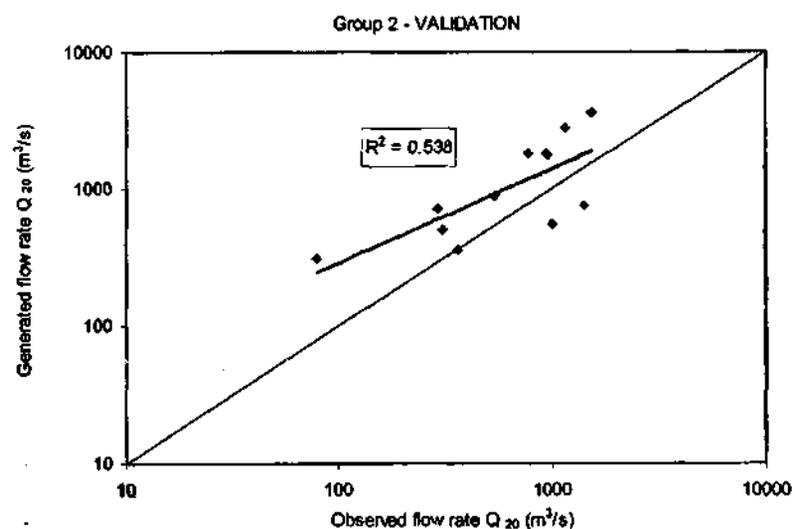


Figure 5
Generated flow rates (based on the regression of Fig. 4) vs. observed flow rates for the 20-year return period - validation set

are based on the negligible difference of the R^2 statistic in calibration and validation when additional landscape data are added to the catchment area, implying no significant additional prediction input from these parameters. The results are summarised in Table 2.

What is evident from the table is that the one group of data is quite different from the other; Group 1 is stronger in calibration and validation than Group 2. The grouping was a random choice process and this result is probably due to the small sizes of the groups (respectively 13 and 12 stations). Larger data sets are probably required to eliminate the effects of outliers in the samples. Table 2 also shows that catchment area on its own is a sufficiently good predictor of floods and the addition of landscape data does not improve this by much. This observation is based on the values of R^2 for the validation group. Besides Area, when Relief and then MAP are included, the results are best for validation using Group 2 (0.556 and 0.553 respectively compared to 0.538 for Area only). Conversely, when Group 1 is used for validation, Area alone has

the best R^2 (0.856) followed by the inclusion of *Drainage Density* (0.784) and *MAP* (0.770). *Drainage Density* is an area surrogate, so the apparent strength in validation might be due to this fact, whereas *MAP* is independent of area.

Ultimately, designers require an efficient flood formula and the acquisition of landscape data is not easy nor does it seem to provide much help to use a more complicated formula. Thus the use of the RMF (area-based) empirical equations seems justified. However, since this is only a preliminary review, further investigations into the role of landscape in affecting a flood regime is required to help with the understanding of this phenomenon.

Return period of the RMF

The RMF method of flood computation was applied to 57 catchments, where both annual flood peak data were available and where the regional K -values were known from Kovacs (1988). The floods were estimated for Regions 4.6, 5 and 5.2, which nearly cover the entire country (the remaining regions have a small number of recorded floods in their database). The floods were modelled from historical records using the GEV distribution and were plotted coaxially with those that were determined from the RMF, corresponding to the same regions and catchments, against catchment area as the independent variable as shown in Fig. 6. Since the return periods of the modelled floods were known, the return period of the RMF could then be estimated. For this reason, the 50-, 100- and 200-year floods were determined from the statistically modelled floods to determine the return period of the RMF. The results are shown in Figs. 6, 7 and 8 which cover Regions 4.6, 5 and 5.2 respectively.

The 200-, 100- and 50-year observed flood magnitudes are represented by the thin solid line, the thin dashed line and the thin dotted line respectively. These magnitudes were determined from the statistical analysis of observed flood data for the individual catchments using the GEV distribution; a subset of the full data set (used for comparison with landscape analyses) appears in Part 3 of Table A1 in the Appendix. The RMF estimates were then determined from Kovacs (1988) using the Francou-Rodier equation and Kovacs' regional K -values for the corresponding catchments. These are represented by the thick solid lines in Figs. 6, 7 and 8. The 200-, 100- and 50-year flood estimates are plotted coaxially with the RMF estimates for the corresponding

TABLE 2
Results of the step-wise regression

				R^2	
				Group 1: calib. Group 2: valid.	Group 2: calib. Group 1: valid.
Q_{20} vs..	Area	Calibration:	0.856	0.538	
		Validation:	0.538 (3)	0.856 (1)	
	Area and slope	Calibration:	0.869	0.566	
		Validation:	0.534	0.724	
	Area and MAP	Calibration:	0.886	0.552	
		Validation:	0.507	0.770 (3)	
	Area and drainage density	Calibration:	0.872	0.538	
		Validation:	0.531	0.784 (2)	
	Area and relief	Calibration:	0.875	0.644	
		Validation:	0.556 (1)	0.628	
	Area and ruggedness number	Calibration:	0.880	0.593	
		Validation:	0.552	0.659	
	Area, slope and MAP	Calibration:	0.896	0.635	
	Validation:	0.502	0.393		
Area, ruggedness number and MAP	Calibration:	0.920	0.596		
	Validation:	0.523	0.640		
Area, drainage density and MAP	Calibration:	0.887	0.571		
	Validation:	0.488	0.534		
Area, relief and MAP	Calibration:	0.890	0.647		
	Validation:	0.553 (2)	0.605		

(1), (2) and (3): These numbers in parentheses flag the "best" (based on the R^2 statistic) fit to the validation data.

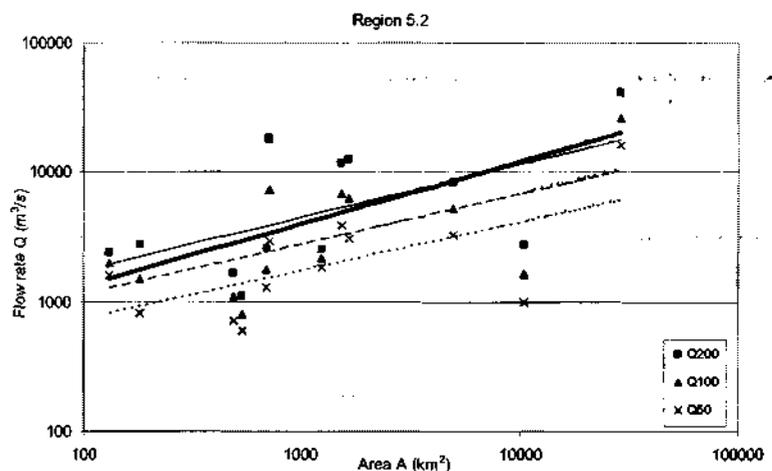


Figure 6
Determination of the return period associated with the RMF for Region 5.2. The bold line is the RMF estimate explained by the Francou-Rodier equation and the thin lines (dotted, dashed and solid) are trend-lines fitted to the 50-, 100-, and 200-year floods estimated from the recorded flows (points) in region 5.2, using the GEV distribution.

catchments in regions 4.6, 5 and 5.2 to examine if a return period can be associated with the RMF.

From Figs. 6, 7 and 8 it is clear that the RMF, when compared to the 50-, 100- and 200-year floods, is closest to the 200-year flood. The trend-line equations, summarised in Table 3, make for interesting reading. The slopes of the corresponding curves and

trend-lines are not equal. In fact the 200-, 100- and 50-year trend-line slopes are slightly flatter than the RMF curves for all the floods in all the regions except one (Q_{50} for Region 4.6); nevertheless, the correspondence is good and provides a starting point for further research to explain the similarities, even if the R^2 values are quite low.

In all three cases, the RMF line and the 200-year trend-line estimated from the fitted GEV distributions almost lie on top of each other and are very nearly parallel, mindful of the contents of Table 3. However, it must be admitted that the trend-lines for the 200-, 100- and 50-year floods have a poor fit and a fair amount of scatter can be observed. In all of Figs. 6, 7 and 8, serious outliers are evident for catchments with areas of about 1 000 km², where the RMF is more likely to be associated with an event of return period of 100 years. On further investigation, it turns out that the problem 200-, 100-, and 50-year GEV flood estimate outliers are skewed by excessively large flood peaks that were observed in a relatively short record (between 20 and 30 years) for those catchments. It is expected that with more data, the effect of the outliers will be diminished. The result is that the plot for Region 5 (Fig. 7) is likely to be closer to the truth than Figs. 6 and 8 as it contains more data. In Fig. 7 the difference between the 200-year flood estimate and the RMF estimate is greater than the other two, but this difference is not excessive and the 200-year estimate and the RMF estimate are of the same magnitude. Based on these results, it is the opinion of the authors that it would be reasonable to assume the RMF to have a return period of the order of 200 years.

Q_T/Q_2 ratios in scaling floods

The Q_T/Q_2 ratios given in Table 1 are based on the premise that flood peaks in South Africa are log-normally distributed with a coefficient of variation equal to 1.3. To determine whether this average relationship is applicable in a design context, each of the maximum flood peaks recorded for the 130 catchments was assigned a return period based on the Weibull Plotting Position and then scaled down to 10- and 50-year events using the Q_T/Q_2 ratios. These flow rates were then plotted against the 10- and 50-year flow rates determined from modelling the historical records by the GEV distribution, to examine if the scaled flow rates are of the same order as the historical ones.

To simplify the analysis, a simple power-law relationship was sought between the ratios Q_T/Q_2 and T shown in Table 1. This short-cut approach was used instead of computing the percentage points of the lognormal distribution and to see its validity, the relationship is shown in Fig. 9, where a power law curve ($Q_T/Q_2 =$

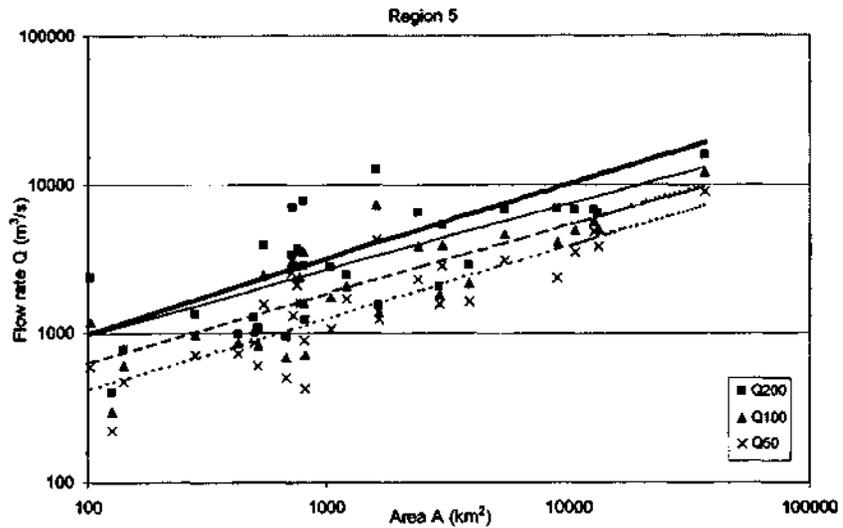


Figure 7
Determination of the return period associated with the RMF for Region 5 (description as per Fig. 6)

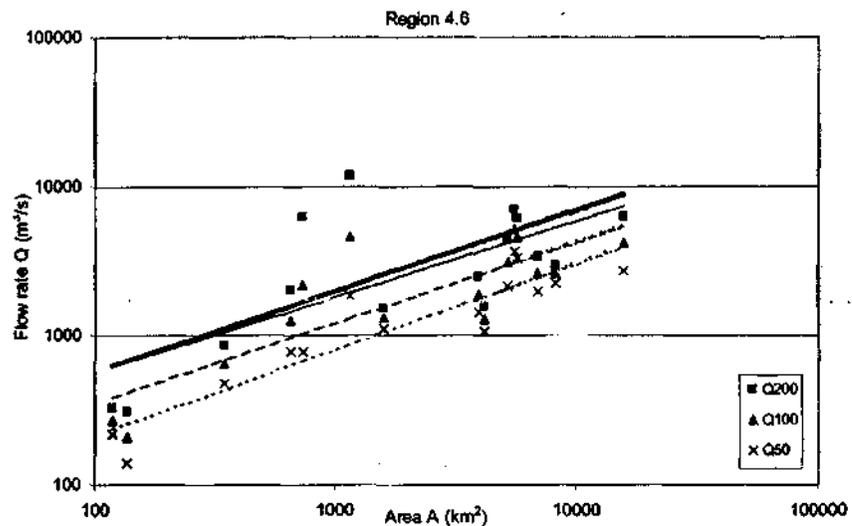


Figure 8
Determination of the return period associated with the RMF for Region 4.6 (description as per Fig. 6)

TABLE 3 Summary of the trend-lines from Figs. 6, 7 and 8				
	Q_{RMF}	Q_{200}	Q_{100}	Q_{50}
Region 5.2	$145A^{0.48}$	$269A^{0.41}$	$191A^{0.39}$	$134A^{0.38}$
Region 5	$100A^{0.5}$	$129A^{0.44}$	$77A^{0.46}$	$45A^{0.48}$
Region 4.6	$48A^{0.54}$	$55A^{0.51}$	$29A^{0.54}$	$20A^{0.58}$

$1.28T^{0.456}$) was fitted in the 10- to 100-year interval. This interval was used firstly because it fitted better ($R^2 = 0.996$) than a power-law curve through all the points and secondly, because all the recorded maxima were observed to lie in this interval, i.e. between 10 and 100 years. The plot of these scaled floods with the corre-

sponding 10- and 50-year floods modelled from historical records using the GEV distribution are shown in Figs 10 and 11.

Figure 10 indicates a reasonably good correspondence for the 10-year event between the GEV modelled flow rate and the scaled flow rate using the fixed lognormal distribution assumption. Fig. 11 indicates an excellent correspondence between the two estimates of flow rates for the 50-year event. The rarer floods are better estimated by the simple ratios than the more frequent ones because the 50-year flood is closer to the average of the record lengths. Even so, there seems to be little bias in the estimates, which might be useful if there is no distributional information. This will have to wait until a regional value of the GEV shape parameter k is obtained - this matter will be addressed below.

The modified rational formula (MRF)

Pegram (2003) introduced the concept of the MRF to enable the rational formula (RF) to be more efficient in flood prediction and more widely used for a variety of catchment sizes. He did this by replacing the rainfall intensity term (the i term in Eq. (5)) with a function that incorporates the median annual maximum rainfall, a scaling function of an extreme value distribution that includes the effect of return period and rainfall duration d . This last term is explicitly replaced by the time of concentration (using the Kirpich formula), which he showed is a function of catchment area, leading to the deduction that the areal reduction factors are nearly constant for a wide range of catchment sizes; the details can be found in Pegram (2003). The combination of these observations allows the MRF to be expressed as a simple function of the median annual maximum rainfall, the reduced variate of an extreme value distribution and catchment area as Eq. (7), repeated here:

$$Q_{MRF} = c \times 0.318 P_{1day,2} [1 + 0.385 y_T] A^{0.556} \tag{7}$$

The formula was introduced as a means of relating various scaling properties of the rainfall-runoff relationship to prompt discussion and further investigation, not as a serious design equation at this stage. Nevertheless, it was intriguing to determine how well it fared in comparison with the other methods described herein. 10-, 20- and 50-year floods based on the MRF were compared with the floods modelled by the GEV distribution from historical records. The values of c and $P_{1day,2}$ (refer to Eq. (7)) used here were based on Pegram's (2003) suggestions which were

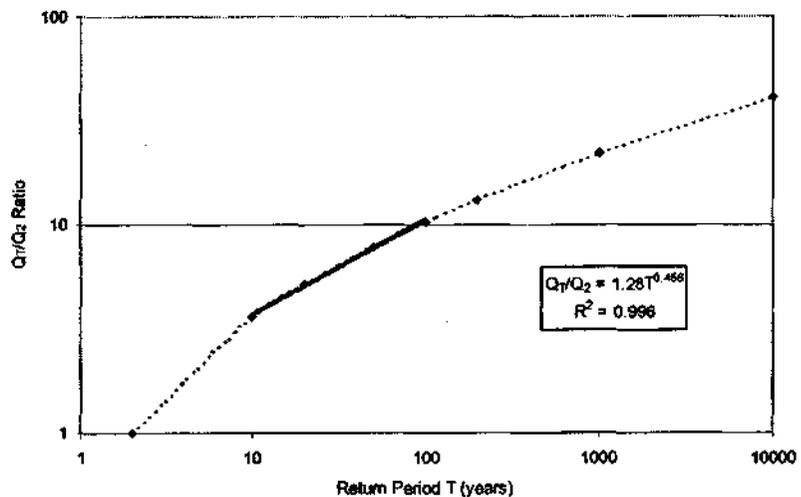


Figure 9
Plot of Q_T/Q_2 ratios vs. return period to determine a simple power-law relationship for the 10- to 100-year recurrence interval T

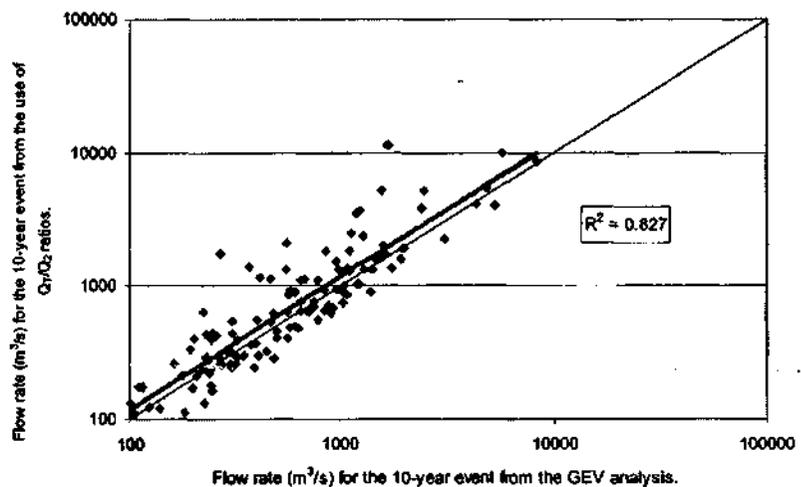


Figure 10
Comparison of Q_{10} flow rates using the Q_T/Q_2 ratios of Table 1 and Fig. 9, compared with those estimated from the GEV model fitted to the observed records

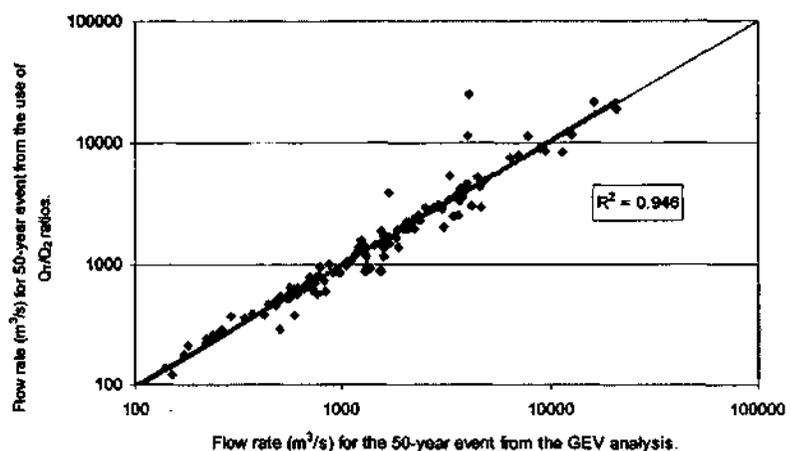


Figure 11
Comparison of Q_{50} flow rates using the Q_T/Q_2 ratios of Table 1 and Fig. 9, compared with those estimated from the GEV model fitted to the observed records

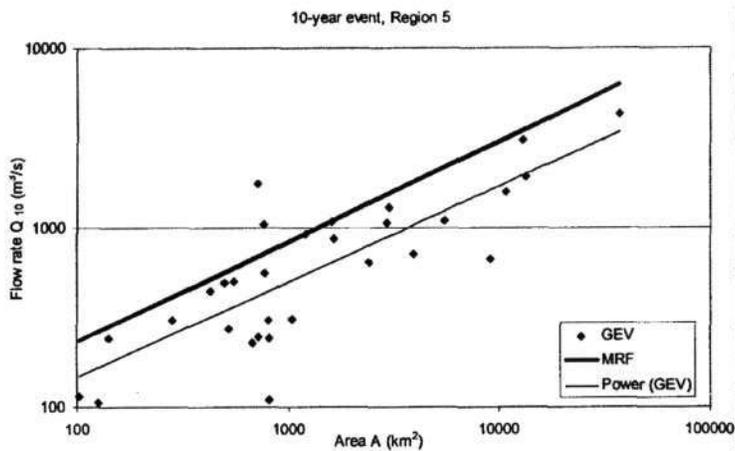


Figure 12

Comparison of the MRF, using $P_{1day,2} = 60$ mm for Region 5 and $c=0.45$ for the 10-year return period and the 10-year flood estimated from the GEV model fitted to the observed records

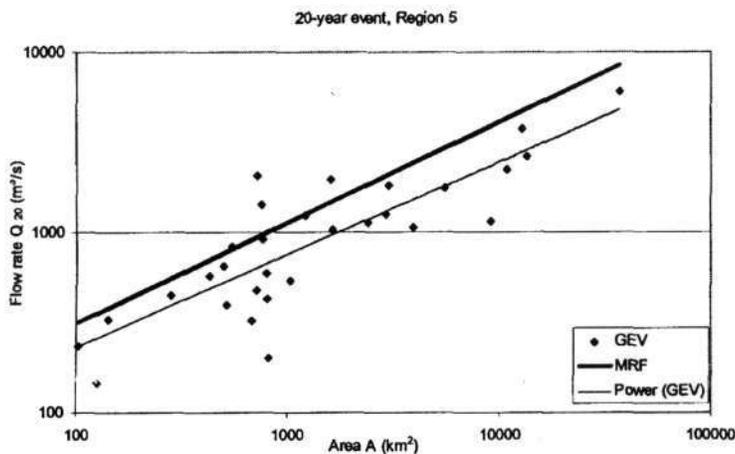


Figure 13

Comparison of the MRF, using $P_{1day,2} = 60$ mm for Region 5 and $c=0.5$ for the 20-year return period and the 20-year flood estimated from the GEV model fitted to the observed records

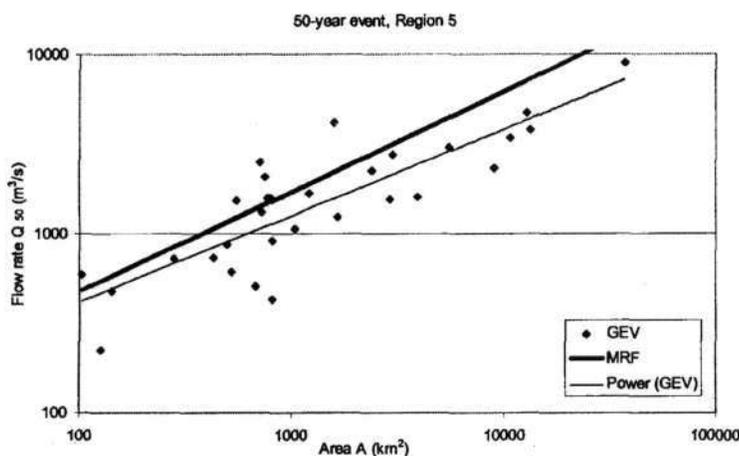


Figure 14

Comparison of the MRF, using $P_{1day,2} = 60$ mm for Region 5 and $c=0.6$ for the 50-year return period and the 50-year flood estimated from the GEV model fitted to the observed records

average values determined for each of Kovacs' K -regions. It is acknowledged that the values have as much as 25% variation within a region, depending on area, size and local conditions, however the aim of this preliminary study is to examine if the MRF's floods are of the same order as those observed and if this alternative is applicable. The regions used were Region 4.6, 5 and 5.2 and the $P_{1day,2}$ values were estimated as 45, 60 and 85mm respectively while values of 0.45, 0.5 and 0.6 were used for c for the 10-, 20 and 50-year events respectively. Values of $P_{1day,2}$ rainfall estimates for each Region K were determined by Pegram (2003) from Adamson's (1981) regional map of median annual maximum 1-day rainfall over South Africa. The value of c could vary from 0.45 to 0.75 (Pegram, 2003), which depends on soil type and land cover. With increasing recurrence interval, the role of rainfall abstractions becomes less as, during larger events, the catchment becomes more saturated and the vegetation gets stripped resulting in an increase in speed and volume of overland flow (AR&R, 2001). Mindful of the fact that large to rare events such as the RMF are likely to have $T > 200$ years which would equate to a high value of c , and that the dominant land type is fairly flat grassland, the subjective estimates of $c=0.45, 0.5$ and 0.6 were made for the 10-, 20-, and 50-year return periods respectively.

The results of Region 5 (which covers the majority of the country) are shown in Figs. 12, 13 and 14. The results for the other two regions are very similar and the trend-line equations for all the regions are summarised in Table 4.

From Figs. 12, 13 and 14 and Table 4 it is evident that the MRF floods and those modelled with the GEV distribution are roughly of the same order of magnitude (although the 10-year GEV trend-line is over-estimated more than the 20- and 50-year events). The MRF floods are generally larger than the GEV floods for the same return period and this is true of all the regions. What is also evident is that the MRF floods model the GEV floods better as the return period increases (larger, less frequent events). This is also true of all the regions. In all figures, a large amount of scatter is observable in the plots of the GEV modelled historical flows. This is especially true for catchments of 1 000 km² area. As the recurrence intervals increase, the scatter becomes less and it would seem that the extreme value distribution is more suited to the 50-year recurrence interval floods.

The results indicate that the MRF computes floods of similar magnitude to those modelled, and the difference might be attributed to the choice of c and $P_{1day,2}$ values whose variation can be of the order of 25%, which Pegram (2003) suggests is probably twice the typical standard error associated with flood peak measurement. The conclusion is that the MRF provides a potential alternative "as a check formula for estimating flood peaks on a wide range of catchment areas for any recurrence interval" (Pegram, 2003). The discrepancy at smaller recur-

TABLE 4
Summary of the trend-line equations from Figs. 12, 13 and 14 for Region 5; and Regions 4.6 and 5.2 which are not shown.

	Region 5.2		Region 5		Region 4.6	
	MRF	GEV	MRF	GEV	MRF	GEV
Q_{10}	25.2A ^{0.558}	49.2A ^{0.357}	17.8A ^{0.558}	12.5A ^{0.534}	13.3A ^{0.558}	2.51A ^{0.679}
Q_{20}	34.0A ^{0.558}	79.6A ^{0.359}	24.0A ^{0.558}	22.1A ^{0.511}	18.0A ^{0.558}	5.78A ^{0.630}
Q_{50}	81.7A ^{0.558}	133.6A ^{0.374}	36.5A ^{0.558}	45.5A ^{0.482}	27.4A ^{0.558}	15.0A ^{0.577}

TABLE 5
Statistics of the GEV k and Kovacs' regional K

Kovacs K	Mean GEV k	Standard Error
5.2	-0.671	0.102
5	-0.477	0.0585
4.6	-0.463	0.113

rence intervals might be attributable to several sources - the choice of coefficients c based on return periods only rather than including more detailed considerations of land type and cover, the use of the median instead of the 10- or 20-year rainfall (also available from maps (Adamson, 1981)), the GEV shape parameters of the rainfall, chosen as a fixed $k = -0.18$, which is far less skew than streamflow which averages $k = -0.5$ (see below). What is encouraging is that the slopes of the lines are not that different. For larger events, the MRF flood estimates are closer to the GEV modelled floods.

Regional GEV distribution values

The parameters of the GEV distribution include a shape parameter k . The relationship between this value and the regional K -value introduced by Kovacs (1988) is now examined to detect if any trend exists. The GEV k was allowed to assume any fitted value above -1.5. The lower limit was placed on the distribution in cases where extreme outliers skewed the extrapolation to unreasonable flood magnitudes. The GEV k , because it was determined by minimising the least squares in the model, was hoped to display a trend with the regional K -values. The GEV k and the regional K were plotted against each other for Regions 4.6, 5 and 5.2. The results are summarised in Table 5 and are shown in Fig. 15.

From Fig. 15, it is evident from the mean that no strong relationship exists. The GEV k -value for each catchment is plotted against the regional K -value for the region in which it is found. The mean GEV k for each region, plus and minus a standard deviation, are also plotted and show that the least deviation is for Region 5, where most of the data are. However, no real conclusion can be attributed to only three points and the lack of catalogued regional K -values for many of the catchments possibly account for this shortfall. The results are inconclusive; however, it would be nice to think that a relationship did exist (as shown by Kjeldsen et al., 2002) and that the modelling of historical floods could be regionalised, perhaps using some regional definition other than the RMF.

Conclusion

It is difficult to place exact values on flood magnitudes and their probabilities of occurring. No one method (empirical, deterministic and probabilistic) can be accepted *a priori* as better than the other, as all methods are approximations and their accuracy is

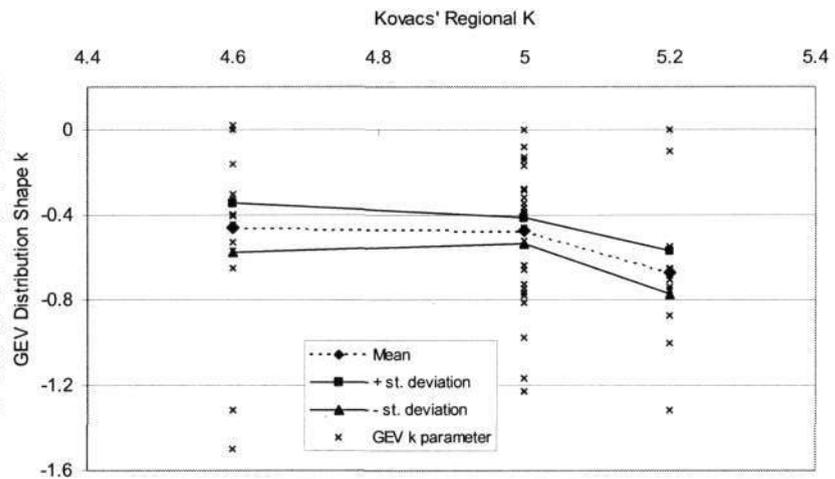


Figure 15
Plot of GEV k (shape parameter) and mean plus and minus a standard deviation, vs. Kovacs' regional K -value for Regions 4.6, 5 and 5.2

relative. However, if one has a sufficiently long period of record, then one can make reasonably accurate predictions on future floods based on what was observed in the past. This premise is based on the assumption that climatic and geological controls in flood production remain the same as when the floods were observed. The database of annual flood peaks for the 130 catchments around South Africa allowed us to utilise the probabilistic method in determining future flood magnitudes that were, in all probability, representative of the truth. This database was used as a reference set to validate the other methods in review.

The addition of landscape data in an attempt to improve flood prediction was not particularly dramatic, nor did it provide much insight into which geomorphometric controls are involved in flood response. The role of landscape in flood production is not in doubt; however, further investigation is required in this regard. Larger sets of landscape data at accurate scales are required to assist in this endeavour. In this preliminary study it was found that landscape data do not make flood prediction more accurate and the area-based RMF empirical equation seems best. If an understanding of the role of landscape in flood production is required then the study should perhaps be based on physical and scaling, rather than statistical relationships.

The RMF method of flood computation is a versatile and easy method to compute the upper limit of floods that a site or region has historically experienced within the observed period of record. In this paper it was found that the return period of the RMF is approximately 200 years and this will no doubt increase the popularity of this method. If the return period is known, the RMF

can now be used as a convenient check method to ascertain the validity of design flood magnitudes. Once the data set used in this paper is augmented by 15 or more years of observed annual floods, refinements are likely to appear.

The use of the Q_T/Q_2 ratios, suggested in TRH 25 (1994), was found to scale observed flood maxima to flow rates with desired return periods reasonably accurately. The Weibull Plotting Position was used to associate a return period with the flood maxima and was based on the flood having been observed within the 10- to 100-year interval. The floods were reduced to 10- and 50-year flow rates through interpolation of the ratios; extrapolation to larger events is yet to be tested.

The use of the MRF proved to be a useful modification of the conventional rational formula and predicted flood magnitudes of similar order to those observed. It has potential to provide an alternative to other flood computation methods and might be used as a check formula for design floods for a wide variety of catchment sizes and return periods, especially where site-specific rainfall data and c -values are used.

The GEV scaling parameter k did not show any particular trend with the regional K -values. Other definitions of regionalisation may allow a stronger relationship; it is known that more arid areas experience floods with higher skew which in turn should give more negatively biased k -values. Our results were inconclusive, but again, more observed flood data should refine this.

The thoughts and ideas presented in this paper are here to serve as preliminary insight into some interesting relationships between some of the flood determination methods employed in South Africa. Admittedly, the flood database used in this investigation is 15 years old, and in many instances, a fair degree of scatter is observable in the flood distributions. Longer records of floods are required to possibly eliminate these anomalies. However, the flood database utilised did provide a consistent foundation to launch this review.

References

- ADAMSON PT (1981) Southern African Storm Rainfall. Tech Rep TR102. Dept. Environ. Affairs, Div. Hydrology, Pretoria.
- ALEXANDER WJR (2002) The Standard Design Flood – Theory and Practice. Dept of Civil Engineering, University of Pretoria, Pretoria.
- AUSTRALIAN RAINFALL AND RUNOFF (2001) *A Guide to Flood Estimation*. Pilgrim DH (editor-in-chief). The Institution of Engineers, Australia.
- DE MICHELE C and SALVADORI G (2002) On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition. *J. Hydrol.* **262** (1-4) 245-258.
- FSR (1975) *Flood Studies Report. Vol. 2 Meteorological Studies*. Natural Environment Research Council, London.
- HIEMSTRA LAV and FRANCIS DM (1979) The Runhydrograph – Theory and Application for Flood Predictions. Water Research Commission, Pretoria.
- HORTON RE (1932) Drainage basin characteristics. *Trans. Am. Geophys. Union.* **13** 350-361.
- HORTON RE (1945) Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology. *Bull. Geol. Soc. Am.* **56** 275-370.
- KJELDSEN TR SMITHERS JC and SCHULZE RE (2002) Regional flood frequency analysis in the KwaZulu-Natal Province, South Africa, using the index-flood method. *J. Hydrol.* **255** (1-4) 194-211.
- KIRPICH ZP (1940) Time of concentration of small agricultural watersheds. *Civ. Eng.* **10** (6) 362.
- KOVACS Z (1988) Regional Maximum Flood Peaks in Southern Africa. Technical Report TR 137, Department of Water Affairs, Pretoria.
- MIDGLEY DC, PITMAN WV and MIDDLETON BJ (1994) Surface Water Resources of South Africa, 1990. WRC Report No. 298/1/94.
- PARAK M (2003) The Flood-Landscape Relationship. Unpublished undergraduate dissertation. School of Civil Engineering, University of KwaZulu-Natal, Durban.
- PEGRAM GGS (2003) Rainfall, rational formula and regional maximum flood – Some scaling links. Keynote Paper. *Aust. J. Water Resour.* **7** (1) 29-39.
- PETRAS V and DU PLESSIS PH (1987) Catalogue of Hydrological Parameters. Flood Studies. Technical Note No. 6, Department of Water Affairs, Pretoria.
- SHREVE RL (1966) Statistical law of stream numbers. *J. Geol.* **74** 17-37.
- SMITHERS JC and SCHULZE RE (2003) Design Rainfall and Flood Estimation in South Africa. WRC Report 1060/1/03, Water Research Commission, South Africa.
- STRAHLER AN (1952) Hypsometric (area-altitude) analysis of erosional topography. *Bull. Geol. Soc. Am.* **63** 1117-1142.
- STRAHLER AN (1964) Quantitative geomorphology of drainage basins and channel networks. In: VT Chow (ed.) *Handbook of Applied Hydrology*. Sec. 4-11. McGraw Hill, New York.
- TRH 25 (1994) *Guidelines for the Hydraulic Design and Maintenance of River Crossings. Vol. 1 2-21 Hydraulics, Hydrology and Ecology*. Pretoria, Department of Transport.
- WESTRAY L (2001) Linking Rational Formula to the Regional Maximum Flood formula to give the Modified Rational Formula (MRF). Unpublished undergraduate dissertation. School of Civil Engineering, University of KwaZuluNatal, Durban.

Appendix

Hydrologic and morphometric parameter definitions (from Parak, 2003):

1. Catchment area: A (km^2)

This is the gross catchment area that is represented by the relevant gauging station.

2. Ineffective area: A_i (km^2)

This includes those areas from which runoff cannot reach the river, for example from pans or depressions.

3. Effective area: A_e (km^2)

This is the part of the catchment area that contributes runoff to the rivers and ultimately the relevant gauging station. It is the difference between the gross catchment area and the ineffective area.

4. Catchment perimeter: P (km)

This is the distance measured along the catchment boundary, i.e. the distance along the watershed boundary.

5. Longest water course: L (km)

This is the distance from the gauging station along the longest watercourse to furthest point on the channel, i.e the start of the permanent streams (fingertip tributary) near the catchment boundary.

6. Maximum elevation above sea level: Z_m (m)

This is the point of maximum height above sea level on the catchment divide (watershed).

7. Elevation of gauging station above sea level: Z_o (m)

This is the elevation of the gauging station above sea level.

8. Catchment relief: R (m)

This is defined as the height difference between the maximum elevation on the watershed and the elevation of the gauging station. It is thus the difference between Z_m and Z_o .

9. Mean river slope: S

The mean slope of the channel or river is computed from the longitudinal profile of the river along the main watercourse from the furthest point along the channel (fingertip tributary) by equating the cut and fill areas (refer to Fig. A1).

10. Shape factor: A/A_c

This is the ratio of the gross catchment area (A) to the area of a circle (A_c) drawn from the longest possible catchment diagonal (refer to Fig. A2).

11. Time of concentration: T_c (hours)

This is the time required for a water particle to travel from the catchment boundary along the longest watercourse (L) to the gauge at the basin outlet, and was computed from Kirpich (1940):

$$T_c = 0.0633 \left(\frac{L}{S} \right)^{0.385} \quad (A1)$$

12. Mean annual precipitation: MAP (mm)

This was determined on 1:250000 scale isohyetal maps.

13. Mean annual runoff: MAR ($10^6 m^3$)

This was obtained from the gauging records.

14. Maximum observed flood peak: Q_{max} (m^3/s)

These are the maximum observed flood peaks recorded at the relevant stations for the length of the representative period.

15. Representative period: N (years)

This is not the return period, but the length of record at the gauges.

16. RMF K -value of region

This is the K -value of the region where the gauging station is found, determined from the Map of Maximum Flood Peak Regions in Southern Africa (Kovacs, 1988). This map divides South Africa into hydrologically homogeneous regions.

17. Total length of all streams in basin: ΣL (km)

This is the sum of the lengths of all the streams and channels which feed the catchment outlet at the gauge. This was determined from a 1:250000-scale map.

18. Strahler basin order

This is the order of the channel, ordered according to the method of Strahler (1952), at the catchment outlet (see Fig. A3(a)). The smallest recognisable channels (fingertip tributaries) are designated order 1. Where two channels of order i join, a channel of order $i+1$ forms downstream. Where a channel of order i and $i+1$ join, the channel downstream assumes the order of the higher channel. These were determined from 1:250000-scale maps.

19. Shreve magnitude

This is the order of the channel, ordered according to the method of Shreve (1966), at the catchment outlet (see Fig. A3(b)). The Shreve magnitude of a catchment stream network reflects the number of fingertip tributaries feeding the catchment outlet. At the

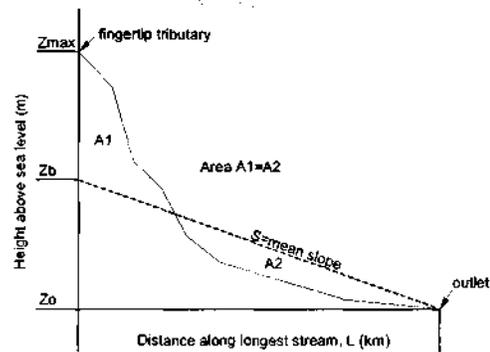


Figure A1

The mean channel slope (after Petras and Du Plessis, 1987)

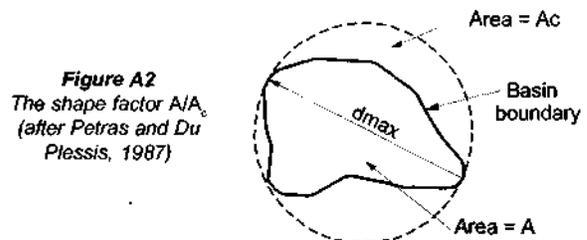


Figure A2
The shape factor A/A_c (after Petras and Du Plessis, 1987)

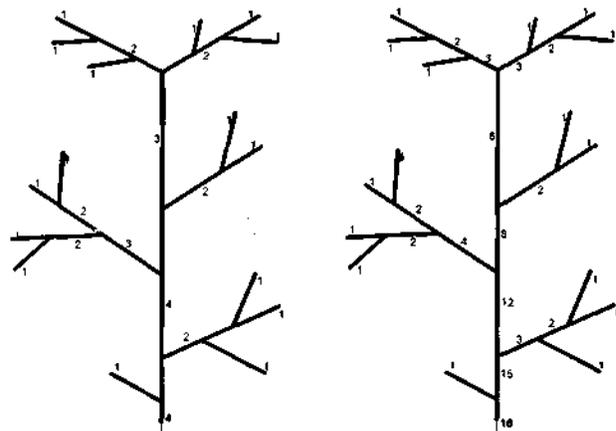


Figure A3

Two methods of stream ordering
(a) Strahler basin order
(b) Shreve magnitude

junction of two streams, the resulting order of the downstream channel is the sum of the orders of the two streams feeding it. These were determined from 1:250000-scale maps.

20. Drainage density DD (km/km^2)

This is the quotient of total stream length within a catchment and the catchment area, determined by dividing ΣL by A .

21. Ruggedness number RN

This is defined as the product of drainage density and catchment relief, i.e. $RN=DD \times R$.

22. Bifurcation Ratio R_b

This was determined by plotting the number of streams in each order of a catchment on a logarithmic scale with the stream order on a linear scale. The slope of the fitted regression line is the bifurcation ratio.

TABLE A1
Hydrological and morphometric data for those catchments in South Africa whose data overlapped/intersected (Part 1)

Gauge number	River	Petras and Du Plessis (1987)										
		Catchment area (km ²)	Ineffective area (km ²)	Effective area (km ²)	Catchment perimeter (km)	Longest water course (km)	Catchment relief (m)	Mean river slope	Shape Factor A/A _c	Time of conc. (hours)	Mean annual precipitation (mm)	Mean annual runoff (x 10 ⁶ m ³)
A2H003	HEX	495	6	489	103.0	36.0	616	0.00917	0.49	6.4	705	15.0
A2H004	PLAT	137	0	137	57.0	22.0	480	0.01320	0.68	3.8	698	7.5
A2H007	APIES	142	0	142	51.0	16.0	289	0.01220	0.42	3.1	770	3.6
B1H001	OLIFANTS	3 989	0	3 989	263.0	172.0	395	0.00155	0.66	42.0	700	127.0
B2H001	BRONKHORST	1 594	0	1 594	183.0	78.3	308	0.00288	0.39	18.1	700	50.0
C1H002	KLIP	4 152	0	4 152	330.0	181.0	820	0.00132	0.56	47.0	785	325.0
C4H001	GROOT VET	5 590	0	5 590	350.0	127.0	645	0.00151	0.72	34.0	218	218.0
C5H007	RENOSTER	348	0	348	94.0	39.8	305	0.00332	0.37	10.2	550	8.7
C6H001	VALS	5 674	50	5 624	414.0	260.0	940	0.00119	0.29	64.0	629	209.0
C8H001	WILGE	15 673	0	15 673	530.0	377.0	1 790	0.00047	0.66	122.0	756	717.0
D1H001	WONDER	2 397	0	2 397	233.0	99.4	831	0.00362	0.52	19.9	488	48.0
H1H006	BREE	753	0	753	132.0	50.0	1 768	0.01130	0.66	7.6	1 040	271.0
H2H003	HEX	718	0	718	152.0	60.8	1 966	0.01130	0.43	8.8	939	162.0
L1H001	SOUT	3 938	0	3 938	270.0	124.0	945	0.00397	0.65	23.0	222	26.0
N3H001	VOEL	1 597	0	1 597	195.0	129.0	1 717	0.00485	0.32	22.0	368	39.0
Q1H001	GROOT VIS	9 091	10	9 081	474.0	107.0	1 449	0.00545	0.51	18.0	370	950.0
Q6H001	BAVIAANS	694	0	694	128.0	59.5	1 254	0.01067	0.49	8.9	620	22.0
R2H005	BUFFALLO	411	0	411	100.0	45.0	939	0.00635	0.57	8.7	880	43.0
S3H002	KLAAS SMITS	796	0	796	141.0	52.0	883	0.00660	0.62	9.6	505	21.0
T3H004	MZIMHLAVA	1 029	0	1 029	157.0	98.0	1 027	0.00411	0.32	18.8	885	106.0
T5H004	MZIMKULU	545	0	545	109.0	42.5	1 832	0.13640	0.55	6.2	1 200	236.0
V3H005	SLANG	676	0	676	125.0	73.8	594	0.00335	0.60	16.4	955	139.0
V6H002	TUGELA	12 862	0	12 862	492.0	297.0	2 571	0.00318	0.65	48.0	920	2 946.0
W5H006	SWARTWATER	180	0	180	60.5	29.0	469	0.01096	0.61	5.0	955	35.0
X1H001	KOMATI	5 499	0	5 499	378.0	207.0	1 092	0.00504	0.48	31.0	890	623.0

TABLE A1
Hydrological and morphometric data for those catchments in South Africa whose data overlapped/intersected (Part 2)

Gauge number	River	Kovacs (1988)					Extracted from Midgley et al. (1994)					
		Max obs. flood peak (m ³ /s)	Represent. period (years)	Regional K-value	Q _{REF} (m ³ /s)	Total stream length (km)	Strahler order	Shreve magnitude	Drainagedensity (km/km ²)	Ruggedness number	Bifurcation ratio	
A2H003	HEX	709	25	5	2 225	144.9	3	12	0.296	0.183	1.55	
A2H004	PLAT	126	37	4.6	682	46.8	2	2	0.342	0.164	2	
A2H007	APIES	454	79	5	1 192	24.1	2	2	0.169	0.49	2	
B1H001	OLIFANTS	1 330	47	4.6	4 212	1 003.2	5	53	0.251	0.0993	1.62	
B2H001	BRONKHORST	1 030	47	4.6	2 567	352.0	3	21	0.221	0.068	1.62	
C1H002	KLIP	1 220	69	4.6	4 304	1 286.7	4	96	0.31	0.254	2.1	
C4H001	GROOT VET	4 380	65	4.5	5 054	1 608.0	4	129	0.288	0.186	1.82	
C5H007	RENOSTER	800	65	4.6	1 128	99.0	2	9	0.284	0.0868	1.29	
C6H001	VALS	2 560	75	4.6	5 094	1 422.4	4	91	0.253	0.238	1.57	
C8H001	WILGE	3 120	74	4.6	8 818	5 243.0	5	415	0.335	0.599	1.86	
D1H001	WONDER	2 690	76	5	4 896	349.1	2	2	0.146	0.121	2	
H111006	BREE	1 140	66	5	2 744	317.1	3	33	0.421	0.745	2.03	
H2H003	HEX	737	38	5	2 680	346.5	4	41	0.483	0.949	2.03	
L1H001	SOUT	2 240	55	5	6 275	1 133.7	5	74	0.288	0.272	1.79	
N3H001	VOEL	1 980	19	5	3 996	530.5	4	37	0.332	0.57	1.48	
Q1H001	GROOT VIS	2 640	70	5	9 535	1 940.4	5	117	0.214	0.31	2.56	
Q6H001	BAVIAANS	895	19	5.2	3 341	99.8	2	7	0.144	0.18	1.4	
R2H005	BUFFALLO	1 050	41	5.4	3 329	84.7	2	3	0.206	0.194	1.5	
S3H002	KLAAS SMITS	906	25	5	2 821	73.7	2	2	0.0926	0.0818	2	
T3H004	MZIMHLAVA	759	41	5	3 208	435.0	4	47	0.423	0.434	2.01	
T5H004	MZIMKULU	651	39	5	2 335	274.6	3	35	0.504	0.923	1.64	
V3H005	SLANG	400	33	5	2 600	251.4	3	25	0.372	0.221	2.04	
V6H002	TUGELA	4 610	41	4	11 341	5 143.5	6	486	0.4	1.028	1.92	
W5H006	SWARTWATER	600	37	5.2	1 748	71.5	3	8	0.397	0.186	2	
X1H001	KOMATI	3 420	49	5	7 416	1 697.1	4	118	0.309	0.337	1.52	

Hydrological and morphometric data for those catchments in South Africa whose data overlapped/intersected (Part 3)											
Gauge number	River	GEV modelled flow rates from recorded flood data							Flow rates scaled by Q_1/Q_2 ratios in Table 1		
		Length of record (years)	Q_1 (m ³ /s)	Q_{10} (m ³ /s)	Q_{20} (m ³ /s)	Q_{50} (m ³ /s)	Q_{100} (m ³ /s)	Q_{200} (m ³ /s)	Shape parameter k	Q_{10} (m ³ /s)	Q_{50} (m ³ /s)
A2H003	HEX	24	185	499	648	869	1 060	1 273	-0.1692	446	976
A2H004	PLAT	42	9	49	79	140	209	309	-0.5301	67	146
A2H007	APIES	46	78	241	330	476	613	780	-0.2816	171	374
B1H001	OLIFANTS	45	127	651	945	1 437	1 908	2 490	-0.3057	633	1 384
B2H001	BRONKHORST	47	20	608	829	1 111	1 320	1 527	0.0140	490	1 072
C1H002	KLIP	71	164	578	774	1 063	1 310	1 586	-0.1628	489	1 069
C4H001	GROOT VET	24	447	1 619	2 361	3 704	5 100	6 942	-0.4037	1 804	3 940
C5H007	RENOSTER	38	36	207	308	480	651	866	-0.3407	329	720
C6H001	VALS	61	495	1 494	2 129	3 284	4 488	6 081	-0.4081	988	2 160
C8H001	WILGE	60	233	974	1 546	2 736	4 149	6 243	-0.5705	1 212	2 648
D1H001	WONDER	65	123	643	1 123	2 265	3 802	6 344	-0.7280	1 032	2 256
H1H006	BREE	28	401	1 041	1 429	2 111	2 801	3 688	-0.3684	466	1 019
H2H003	HEX	28	101	248	479	1 321	2 997	6 925	-1.2274	386	843
L1H001	SOUT	41	92	717	1 060	1 624	2 157	2 805	-0.2865	994	2 172
N3H001	VOEL	20	145	1 077	1 976	4 189	7 263	12 491	-0.7680	1 405	3 071
Q1H001	GROOT VIS	58	187	671	1 146	2 331	3 995	6 858	-0.7832	1 051	2 297
Q6H001	BAVIAANS	19	498	710	890	1 290	1 797	2 591	-0.6509	635	1 388
R2H005	BUFFALLO	26	392	832	1 001	1 218	1 382	1 544	-0.0001	531	1 161
S3H002	KLAAS SMITS	29	111	306	594	1 590	3 477	7 711	-1.1648	571	1 247
T3H004	MZIMHLAVA	30	36	309	541	1 062	1 726	2 772	-0.6588	384	839
T5H004	MZIMKULU	28	106	500	829	1 552	2 456	3 860	-0.6364	337	736
V3H005	SLANG	34	72	228	327	507	693	940	-0.4053	223	487
V6H002	TUGELA	45	1 563	3 071	3 767	4 787	5 649	6 601	-0.1484	2 333	5 097
W5H006	SWARTWATER	33	37	193	363	821	1 511	2 777	-0.8742	318	694
X1H001	KOMATI	76	205	1 104	1 758	3 066	4 558	6 695	-0.5213	1 598	3 492

The rational formula from the runhydrograph

Mohamed Parak and Geoffrey GS Pegram*

Civil Engineering Programme, University of KwaZulu-Natal, Durban, 4041, South Africa

Abstract

The rational formula is possibly the simplest flood estimation technique available using rainfall-runoff relationships. In spite of the many criticisms regarding its over-simplification of the processes of rainfall conversion into runoff, it remains possibly the most widely used method for estimating peak flood flows for urban drainage systems and small (<100 km²) rural catchments. However, as a result of the criticisms, the formula carries with it many cautions. One such caution regards the determination of the formula's runoff coefficient c , which is seen as the main difficulty in the design application of the formula. Mindful of this, it was decided to investigate the calibration of this coefficient, on past flood peak and flood volume pairs for a number of catchments in South Africa. To this end the "data set" of runhydrographs, which describe the characteristic peak and volume discharges of a catchment for a given recurrence interval, was used to calibrate the coefficients for selected catchments and to explore the assumptions underpinning this simple model. This article describes the methods employed in achieving this as well as a discussion of the results.

Keywords: design flood estimation, probabilistic rational formula, runhydrograph, calibration of runoff coefficients

Introduction

The rational formula is perhaps the best known and most widely used method for the determination of peak flood flows from rainfall events. It has survived numerous criticisms regarding its over-simplification of the complex hydrological processes of flood production but nonetheless is possibly the most favoured method used by practitioners for peak flood estimation. The rational formula owes its popularity to the fact that it is easy to understand and simple to use. The peak flood flow due to a rainfall event on a catchment, determined from the rational formula, is expressed (in SI units) as:

$$Q_{RF} = ciA/3.6 \quad (1)$$

where:

- Q_{RF} is the flood peak in m³/s
- c is the runoff coefficient, which is (in the traditional deterministic approach) defined as the proportion of precipitation that contributes to runoff
- i is the storm rainfall intensity in mm/h
- A is the catchment area in km².

The criticisms concerning the rational formula in the above form are not unfounded and the use of this method carries valid cautions that are based on the following assumptions built into the formula (which are not always explicit in its presentation):

- The maximum rate of runoff from a catchment is achieved when the duration of rainfall is equal to the time of concentration (T_c) of the catchment, which is defined as the time taken for the outflow from a catchment to reach near equilibrium due to a storm uniformly spread in space and time,

- The spatial and temporal characteristics of rainfall are consequently ignored and the storm rainfall, as input into the formula, is assumed to be a rectangular pulse of duration T_c deposited in lumped form on the catchment (i.e. there is no routing component implicit in the formula).

As a consequence, the rational formula was previously limited in its application to small catchments (<15 km² in South Africa (HRU, 1972)) and was only to be used as a check method (it was not to be used in isolation). It was further noted that sound engineering experience and judgment was required for its use. However, work that has since been done, locally by Alexander (2002) and Pegram (2003), and abroad in Australia (Institute of Engineers Australia, 1987), has shown that these cautions were too timid and its use may well be extended beyond small catchments.

For the estimation of design floods, a probabilistic approach to the rational formula is needed, where the variables c and i (the runoff coefficient and rainfall intensity respectively) of the formula are associated with a probability of exceedance. A probabilistic approach is different to a deterministic approach (which is the form shown in Eq. (1)), as it does not involve the representation of a historic event. As opposed to the latter case, no unique combination of rainfall and catchment conditions (such as storm patterns, ground cover conditions, antecedent moisture conditions, etc) exist to reproduce the design flood. In a probabilistic approach, the rational formula is used to estimate, for a given probability of exceedance, the magnitude of the peak discharge from a site; this peak would be equivalent to a discharge estimated from a frequency analysis of flood records if a long and representative record were available at that site.

Pilgrim and Cordery (1993) stated that the design situation is exactly suited to the probabilistic approach of the rational formula and has little similarity with the deterministic rational formula, so that the criticisms associated with the deterministic approach are not necessarily valid for the probabilistic design case. Alexander (1990) stated that as the catchment size increases the value

* To whom all correspondence should be addressed.

☎ +2731 260-3057; fax: +2731 260-1411;

e-mail: pegram@ukzn.ac.za

Received 18 March 2005; accepted in revised form 30 January 2006.

of c becomes more probabilistic than deterministic in its derivation. The probabilistic approach to the rational formula has the same form as Eq. (1) but is defined more specifically as:

$$Q_{(T)} = c_{(T)} i_{(T_c, T)} A / 3.6 \quad (2)$$

where:

$Q_{(T)}$ is the flood peak in m^3/s of recurrence interval (RI) T -years

$c_{(T)}$ is the runoff coefficient for a T -year event

$i_{(T_c, T)}$ is the T -year storm rainfall intensity in mm/h of duration equal to the time of concentration T_c (h) of the catchment

A is the catchment area in km^2 .

In this approach, the value of $c_{(T)}$ purports to transform a T -year design storm $i_{(T_c, T)}$ of duration T_c into a T -year flood peak $Q_{(T)}$ for a catchment of area A . The variable $i_{(T_c, T)}$ can be determined, for a particular site, from suitable intensity-duration-frequency (IDF) relationships of design storms. However, the estimation of the runoff coefficient $c_{(T)}$ remains the main source of uncertainty in the probabilistic application of the rational formula. It is the least precise variable of the rational formula, in spite of it being bounded in the interval (0; 1), and suggests that a fixed ratio of peak runoff rate to rainfall rate exists for the site, which in reality is not the case (Chow et al., 1988: 497). It is this characteristic (the estimation of the design runoff coefficient) of the rational formula that forms the main focus of this article. To this end, this article investigates the calibration of the runoff coefficient, on past flood peak and flood volume pairs for a number of catchments in South Africa, to assist with its determination. The calibration of runoff coefficients on past floods is also the practice that was adopted in Australia (Institute of Engineers Australia, 1987). It was shown in Australia that the use of calibrated coefficients in a probabilistic approach to the rational formula could consistently provide flood estimates for catchments up to $250 km^2$. In this research, the "data set" of runhydrographs produced by Hiemstra and Francis (1979) was used to calibrate the coefficients in order to investigate this for the selected catchments.

In South African practice, the idea of calibrating the rational formula's runoff coefficient is not new. Alexander (2002) proposed a new standardised regional flood estimation technique called the standard design flood (SDF). This method is essentially a probabilistic approach to the rational formula, as advocated by Alexander (1990), utilising calibrated runoff coefficients. The SDF method is based on the calibration of the runoff coefficient against design floods determined from a frequency analysis, using the LOG-Pearson-III (LPIII) distribution, of recorded events from a number of catchments in South Africa. According to Alexander (2002), the SDF can be applied to all sizes of catchments in South Africa, ranging in size from $10 km^2$ to $40\,000 km^2$. Alexander has also suggested a standard design hydrograph for the SDF with a fixed triangular shape that has a rising limb equal to the time of concentration of the catchment T_c and a falling limb equal to $2T_c$, i.e. an effective time base-length of $3T_c$. This idealised hydrograph is the same as that proposed by Rooseboom et al. (1981) where, in this instance, it is noted that the runoff volume is greater than the proportionate part of the storm rainfall that runs off during the time of concentration. In an independent test, the average ratio of Alexander's 50-year SDF flood peak to the 50-year LPIII flood peak was found to be approximately 210% (Görgens, 2002). Alexander's method was designed to be purposefully conservative and he states that the over-estimates fall within the range of uncertainties associated

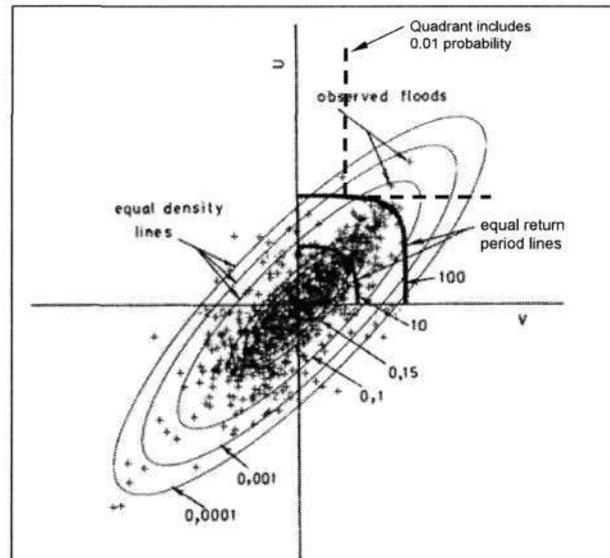


Figure 1

A standard bivariate normal probability density function, with a cross correlation coefficient of 0.85, plotted with log-transformed observed flood peak-volume pairs in probability space (from Hiemstra and Francis, 1979: 14). The bold lines in the positive quadrant are the 10- and 100-year return period joint-exceedance contours. The dashed lines include a quadrant to the upper right which, on average, will include 1% of the observations.

within all design flood procedures. However, Görgens (2002) states that although the cost and implications associated with a conscious over-design in terms of a bridge/culvert is relatively minor, by contrast it is not acceptable for dam spillway design, where the cost of the spillway is a significant component of the total dam cost. An average over-estimate of 200% might render some projects infeasible. As such, Görgens recommends that the SDF should be seen as a conservative approach similar to that of the regional maximum flood (RMF) method.

Conscious of this, it was thought that where this investigation would add value would be in the calibration of the runoff coefficient on past flood peak and volume pairs, as offered by the runhydrograph method. Thus, it was hoped that this would yield coefficients that could, in a design situation, describe a complete design flood hydrograph (peak, volume and time base-length). The following sections describe and explain the theory behind the runhydrograph method, the methods employed in this investigation and the results achieved.

The runhydrograph

The runhydrograph method (Hiemstra and Francis, 1979) summarises, for a given catchment, the family of characteristic peak and volume discharges for a given recurrence interval. These hydrographs were based on the frequency analyses of all rare hydrographs (which were carefully screened for reliability) in a continuous stream flow record and, as such, are independent of rainfall input and catchment characteristics. This set of statistics was thus a valid data set against which to calibrate the runoff coefficient towards a probabilistic approach of the rational formula.

The runhydrograph method was developed by Hiemstra and Francis (1979) (in the sequel referred to as H&F) and was

based on earlier work by Hiemstra (1972; 1973; and 1974), Hiemstra et al. (1976) and Francis (1979). It is based on the joint probability analyses of same-event flood peak and flood volume pairs of recorded data from 43 catchments around South Africa (see Table A1 in the Appendix). H&F discovered that the natural logarithms of the flood peak and its corresponding volume were approximately normally distributed and well correlated, with a cross-correlation coefficient with mean 0.78 and standard deviation 0.12 (a relatively narrow range) whose mode is 0.85. Fig. 1 shows the natural logarithms of the recorded flood peak and volume pairs plotted together with the contours of equal probability density of a standardised bivariate normal probability density function (with a cross-correlation coefficient of 0.85). Also shown in Fig. 1 (in the positive quadrant) are 10- and 100-year return period exceedance probability contours (bold lines). The dashed lines intersecting on the 100-year exceedance contour define an area in the plane whose probability density integrates to 0.01. Thus, on average, 1% of the observations will lie within this area, and within other areas defined similarly on the 100-year contour.

The contours describe the joint probability of flood peak and flood volume exceedance and are able to produce “families” of hydrographs (peak-volume pairs) of equal probability of jointly being exceeded, but of varying shape. These families can range from the marginal peak (associated with any volume), to the “most likely” joint peak and volume pair through to the marginal volume, each with an equal probability of joint exceedance. However, it can be seen from Fig. 1 that the plotted peak-volume pairs cluster around the 45° line in an elliptical shape. If the cross-correlation coefficient approaches unity, the minor axis of the ellipse reduces to zero. Thus, although more than one combination of a peak-volume pair exists that has the same probability of jointly being exceeded, the most likely (modal) pair will be found at the intersection of the 45° line on the exceedance contour, the point where the probability density is highest.

Figure 2 shows the application of the runhydrograph method for design flood peak and volume estimation for a cross-correlation coefficient of 0.85. The listed numbers on the top right of Fig. 3 are the standardised ordinates of the peak-volume exceedance contours for the selected recurrence intervals. They describe the joint exceedance of the most likely peak-volume pair (corresponding to line #1) through to the exceedance of the marginal peak (corresponding to the vertical axis to the left of line #6). It is unlikely that a peak-volume pair will occur on lines

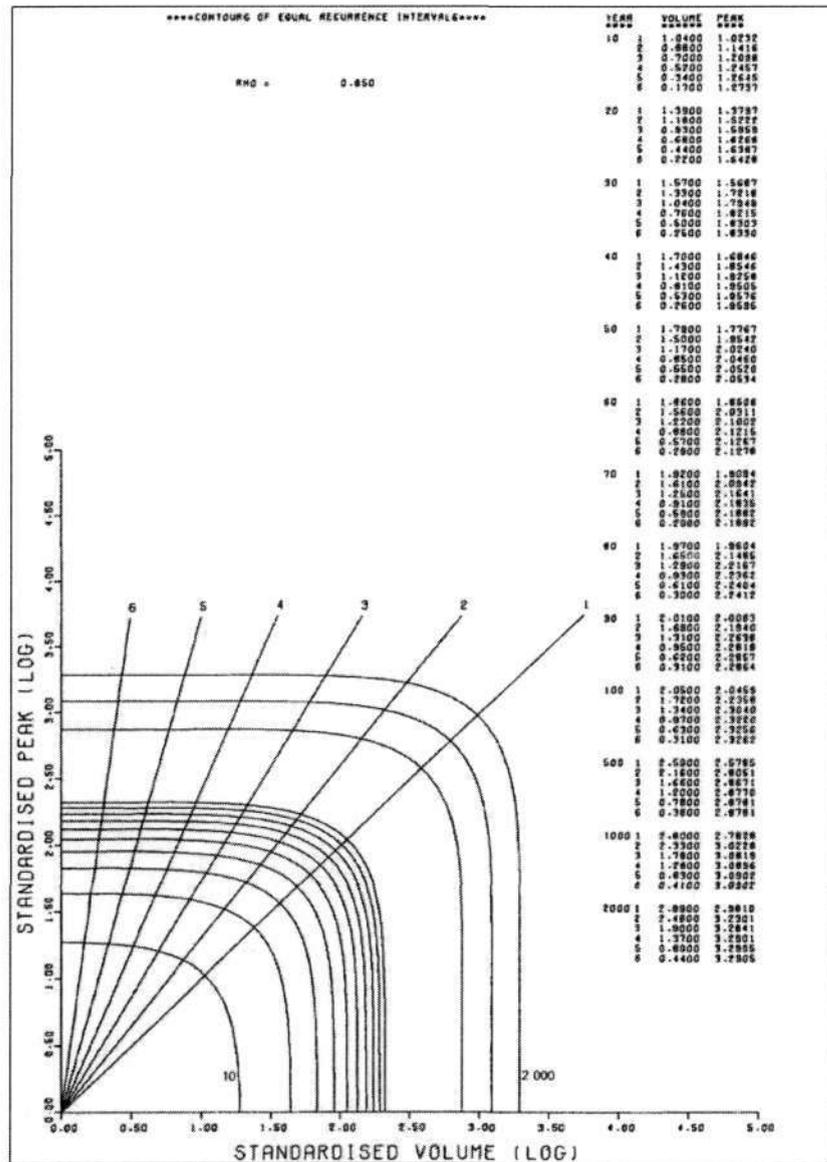


Figure 2 Joint flood peak and flood volume exceedance contours, in probability space for a peak-volume cross-correlation coefficient of 0.85 (from Hiemstra and Francis, 1979)

4, 5 and 6 for this relatively high correlation, and thus for the purposes of this investigation the modal peak-volume pair was chosen in order to limit the number of variables.

In passing, it is interesting to note that this idea of describing hydrographs with a joint probability of exceedance of peak and volume, has surfaced again more recently to be exploited in the evaluation of dam safety (De Michele et al., 2005).

Method and results

The methods employed in this investigation were typical of those used in the derivation of a probabilistic rational formula utilising calibrated coefficients, of which an explanation follows which is adapted from Pilgrim and Cordery (1993) (quotes appearing in italics):

- Where a set of long and reliable record of flood data from a particular catchment exists, a frequency analysis should be carried out on the observed data to determine design values of flood peaks for a range of recurrence intervals. In this study, flood peak and volume pairs (Q_T in m^3/s and V_T in m^3 respectively) for the 'most likely' runhydrograph was computed for each of the selected catchments for RIs of 10-, 20-, 50-, 100- and 200-years. These appear in Tables A2 to A6 in Appendix A. As a result of this, values of B , the time base-length of the triangular approximated hydrographs, were also computed.
- A design formula for the calculation of time of concentration T_c must be selected and used consistently throughout the derivation and use of this method. In this study the Kirpich (1940) formula was used, following the lead of Petras and Du Plessis (1987):

$$T_c = 0.0633[L^2/S]^{0.385} \quad (3)$$

where:

T_c is the catchment's time of concentration (in hours)
 L is the length (in km) of the longest water course
 S is the slope of the longest water course.

- Design rainfall intensities, $i_{(T_c, T)}$, for the corresponding time of concentration of the catchment and recurrence interval should be determined from a suitable Intensity-Duration-Frequency (IDF) database. These were determined from Smithers and Schulze's (2002) design rainfall data-base for South Africa. These data appear in the same tables in the Appendix. (A computer program with a graphical user interface has been developed to obtain design rainfall depths for any location in South Africa from this database. The software may be downloaded from the following website: <http://www.beeh.ump.ac.za/hydrorisk/> and follow the "Design Rainfall" option).
- From these data, values of $c_{(T)}$ can be back calculated by the following equation (where the variables are as defined for Eq. (2)):

$$c_{(T)} = \frac{3.6 \cdot Q_T}{i_{(T, T)} \cdot A} \quad (4)$$

This data also appears in the same tables in the Appendix.

- These calibrated values of $c_{(T)}$ can then be regressed on any physical characteristic of the catchment. In order to validate the calibrated coefficients at untested sites, regional parameters with which to relate $c_{(T)}$ with RI were sought. However, it was noted by Pilgrim and Cordery (1993) that the probabilistic runoff coefficients determined for Australia did not show much sensitivity to physical characteristics of a catchment.

It is important to note that the values of $c_{(T)}$ obtained in this manner are conditioned on the use of a consistent formula for the calculation of T_c and a consistent database for the derivation of the IDF rainfall relationships. A detailed explanation of each of the steps listed above and the results of each exercise are given in the following subsections.

The streamflow database

The 43 catchments used by H&F in their study were based on the Department of Water Affairs and Forestry's drainage region

delineations. These and their derived statistics are listed in Table A1 in the Appendix. As a point of departure, runhydrograph data from H&F were combined with catchment parameters from Petras and Du Plessis (1987), namely area (A) and time of concentration (T_c - based on Kirpich's formula). The number of catchments from the H&F database, for which T_c values were available from the Petras and Du Plessis catalogue, reduced the number of available catchments for calibration to 29. These are listed in Table A2 in the Appendix and formed the core data set on which the rational formula calculations were performed.

The rainfall database

For each of the 29 catchments, a number of locations (depending on the size of the catchment) along the main watercourse within the catchment were chosen for which design rainfall estimates were obtained from Smithers and Schulze (2002). The output from this rainfall database provides point rainfall depths (in mm) for durations ranging from 5 minutes to 7 days and for return periods ranging from 2 to 200 years at a spatial resolution of 1 arc minute in South Africa. The mean depth for each catchment was computed and thereafter the intensity, duration and frequency (IDF) relationships were computed by fitting a simple power-law function of storm duration to the mean rainfall depths. For the selected recurrence intervals, these took the form of: P (rainfall depth in mm) = ad^b and i (rainfall intensity in mm/hr) = ad^c , where d is the storm duration in hours and a , b and c (which equals $b-1$) are the fitted power-law parameters. The mean intensity, corresponding to the time of concentration T_c , was calculated from the IDF relationships for the 10-, 20-, 50-, 100- and 200-year recurrence intervals for each catchment. The parameters fitted to the rainfall duration, for the selected recurrence intervals, are listed in the Appendix (Tables A2 to A8). It was found that rainfall depth scaled, on average, to the power of 0.238 of rainfall duration and thus rainfall intensity to the power of -0.762 of rainfall duration with a standard deviation of 0.0419.

Area reduction factors (ARFs) were not used in this study to scale the point rainfall depths into average depths over the catchment. Instead simple averages of a few representative points along the longest watercourse within the catchment, determined from the Smithers and Schulze (2002) database, were used to account for the variation in precipitation with position and altitude for large catchments. ARFs were deemed not necessary based on the findings of Pegram (2003), of which a summary is presented here. He investigated the scaling properties of rainfall in South Africa and found that they could be expressed as a function of three factors: the median one-day rainfall (which is a function of location), a function of return period (the reduced variate of the general extreme value (GEV) distribution) and a function of duration. He used this finding to modify the intensity expression of the rational formula. The storm duration used by Pegram was the catchment's time of concentration T_c , as in this study, from the Kirpich (1940) formula. When this duration T_c was plotted against catchment area, it was found that the points clustered about a curve to which a power-law relationship could be fitted. This was superimposed on the area reduction factor (ARF) diagram, published in the *Flood Studies Report* (FSR, 1975). He found that the $Area \sim T_c$ curve yielded an almost constant ARF value of 87% across the FSR curve. The implication of this is that, as long as the precipitation intensity used in the rational formula corresponds to the time of concentration of the catchment, the point rainfall is automatically scaled by a constant ARF. It is likely

that the FSR's ARF curves over-estimate the relationship in South Africa, but the degree is likely to be a matter of climate (Pegram and Parak, 2004), so it is also likely that the scaling behaviour will be maintained. However, the reduction factor would automatically be absorbed into the fitted $c_{(T)}$ -values. The first thing to note then is that because c is explicitly a function of T_c , it is therefore implicitly independent of the ARF.

Calibration of the runoff coefficients (c_m)

The next thing to explore was the dependency of c on the flood regime of catchments of various sizes and locations. The first task was to relate c to the peaks of each catchment for varying recurrence interval, T . c -values were fitted to the flood peak of the calculated modal runhydrograph at each site, using the parameters for that site as estimated by H&F.

The summary of results from the calibration of the runoff coefficients is shown below in Table 1. They are the 10-, 20-, 50-, 100- and 200-year runoff coefficients for each of the 29 catchments that formed the core data set. These data and results are drawn from the Appendix (Tables A2 to A6).

Coefficients from 6 of the 29 catchments (marked with an asterisk in Table 1) produced results that did not increase in

magnitude with recurrence interval. As mentioned in Alexander (1990), an increase in c with recurrence interval is necessary to accommodate the known effects which also increase with rainfall intensity but are not accounted for in the formula's calculation process. The main effect, requiring this increase of c with recurrence interval, is that the catchment is likely to be more saturated at the start of a storm with a longer recurrence interval (Rooseboom et al., 1981). This initial saturation caused by pre-event rainfall is the main reason why one can expect to obtain a higher percentage runoff with an increase in the recurrence interval of an event. Alexander (2002) states that in many of the destructive events observed, severe rainfall events were often preceded by above-normal seasonal rainfall.

The calibrated values of c (values of $c_{(T)}$ for all 29 catchments) were coaxially plotted with c -values from Chow et al. (1988: 498) against recurrence interval. This relationship is shown in Fig. 3, where the coefficients from Chow et al. (1988) correspond to the "flat" slopes type (i.e. for ground slopes between 0 and 2%, since all the test catchments in this calibration exercise had slopes of less than 2%) and are for the three "undeveloped" (rural) coverage types (i.e. cultivated land, pasture/range and forest/woodland). These values are shown in Table 2 and were determined for small rural catchments (i.e. less than 100 km²) of Austin, Texas (USA).

Num.	Station	River	Latitude (degrees decimal)	Longitude (degrees decimal)	Catchment area (km ²)	Time of concentration T_c (hours)	Calibrated c -coefficients				
							10-year	20-year	50-year	100-year	200-year
1	A2M03	Hex	25.77	27.28	494	6.4	0.301	0.303	0.304	0.305	0.306
2	A2M12	Krokodil	25.82	27.92	2 586	18	0.089	0.093	0.095	0.097	0.098
3	A3M01	Klein Marico	25.53	26.10	1 002	8.7	0.084	0.092	0.104	0.113	0.123
4	B2M01	Bronkhorstspuit	25.80	28.77	1 585	18.1	0.210	0.228	0.244	0.254	0.262
5	B4M03	Steelpoort	25.02	29.53	2 271	19.6	0.091	0.102	0.112	0.125	0.135
6*	B7M04	Klaserie	24.55	31.03	130	3.7	0.234	0.233	0.214	0.227	0.224
7	C1M01	Vaal	26.95	29.27	8 254	74	0.396	0.419	0.444	0.460	0.476
8	C4M01	Groot Vet	28.48	26.67	5 504	34	0.368	0.386	0.409	0.425	0.442
9*	C4M02	Vet	27.85	25.90	17 550	111	0.179	0.175	0.170	0.167	0.164
10	C5M03	Modder	29.17	26.58	1 650	18.3	0.419	0.440	0.458	0.469	0.479
11	C5M04	Modder	28.85	26.18	5 012	38	0.528	0.592	0.660	0.706	0.749
12	C5M12	Riet	29.65	25.98	2 383	23	0.218	0.235	0.252	0.264	0.274
13	C5M15	Modder	28.80	26.10	6 545	43	0.280	0.302	0.325	0.341	0.355
14	C7M01	Renoster	27.27	27.18	5 255	57	0.236	0.300	0.379	0.438	0.498
15	D1M05	Oranje	30.03	28.50	10 891	60	0.261	0.266	0.270	0.272	0.274
16*	D5M01	Renoster	31.65	20.62	2 129	27	0.263	0.264	0.264	0.264	0.264
17*	D5M04	Sak	31.65	21.77	5 799	28	0.130	0.128	0.125	0.123	0.121
18	E2M02	Doring	32.50	19.53	5 778	30	0.389	0.420	0.459	0.487	0.516
19	H1M06	Bree	33.42	19.27	754	7.6	0.454	0.457	0.461	0.464	0.468
20*	H1M07	Wit	33.57	19.15	83	2.4	0.814	0.800	0.790	0.787	0.786
21	H7M04	Huis	33.92	20.72	26	2.3	0.278	0.307	0.336	0.353	0.368
22	J2M03	Gamka	33.53	21.65	17 941	42	0.076	0.082	0.090	0.095	0.099
23	J3M04	Olifants	33.48	23.03	4 330	23	0.163	0.180	0.194	0.200	0.205
24	Q1M01	Groot Vis	31.90	25.48	9 150	18	0.089	0.097	0.108	0.116	0.124
25	Q9M10	Groot Vis	33.22	26.87	29 376	108	0.176	0.227	0.282	0.318	0.349
26	Q9M12	Groot Vis	33.10	26.45	23 041	85	0.113	0.133	0.158	0.178	0.198
27*	T3M02	Kinira	30.48	28.62	2 100	26	0.186	0.172	0.156	0.145	0.135
28	W4A03	Pongola	27.42	31.52	5 843	31	0.267	0.278	0.284	0.285	0.284
29	W5M05	Hlelo	26.83	30.73	751	17.8	0.177	0.193	0.212	0.225	0.237

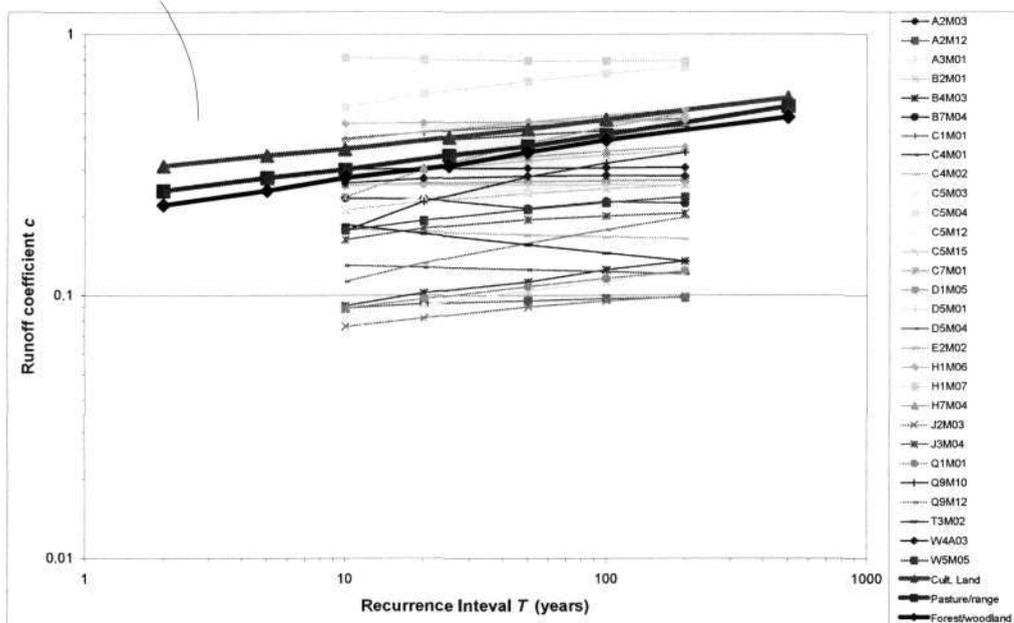


Figure 3
A comparison of the runoff coefficients c from Chow et al. (1988: 498) with those calibrated in this study $c_{(T)}$. The c -values plotted from Chow et al. are shown in thick bold lines and extend from the 2- to 500-year recurrence intervals.

It is evident from Fig. 3 that the $c_{(T)}$ -values obtained from this exercise are spread around those of Chow et al. (1988) but are generally lower in magnitude. The $c_{(T)}$ -values obtained from this exercise range from 0.084 to 0.786, while the values from Chow et al. are between 0.28 and 0.57 (for the recurrence interval range of 10- to 200-years). However, the scatter associated with the latter data set is not known and hence not shown, so it is conjectured that they are curves fitted to the high side of the original data.

Hydrograph time base-length B

The use of flood peak and volume pairs for calibration in this investigation (from the runhydrograph method of H&F (1979)) was thought to have the added advantage in that complete design flood hydrographs could be calculated from these calibrated coefficients. From the flood database computed for the calibration exercise, hydrograph time base-lengths B for each RI were determined for each catchment. Out of interest, they were then expressed as ratios to the catchment's time of concentration T_c for the respective recurrence intervals (which, in terms of the rational formula, is effectively a ratio to the hydrograph's time to peak). The average ratio of B/T_c , for each recurrence interval, was then determined and the results are presented in Table 3 together with their standard deviations. These results exclude three catchments whose area is 130 km² or less as they gave B/T_c ratios in excess of 7. It is noted here that there is an increase of base-length with recurrence interval, which means that the volumes of the floods relative to the peaks, as modelled by the runhydrograph, also increase with T . The figures in the third row of Table 3 show the proportion of floods whose base-

Character of surface	Runoff coefficients c							
	2-year	5-year	10-year	25-year	50-year	100-year	200-year (interpolated)	500-year
Undeveloped								
<i>Cultivated land</i>								
Flat, 0 - 2%	0.31	0.34	0.36	0.40	0.43	0.47	0.51	0.57
Average, 2 - 7%	0.35	0.38	0.41	0.44	0.48	0.51	0.55	0.60
Steep, >7%	0.39	0.42	0.44	0.48	0.51	0.54	0.57	0.61
<i>Pasture/range</i>								
Flat, 0 - 2%	0.25	0.28	0.3	0.34	0.37	0.41	0.46	0.53
Average, 2 - 7%	0.33	0.36	0.38	0.42	0.45	0.49	0.52	0.58
Steep, >7%	0.37	0.4	0.42	0.46	0.49	0.53	0.56	0.60
<i>Forest/woodlands</i>								
Flat, 0 - 2%	0.22	0.25	0.28	0.31	0.35	0.39	0.42	0.48
Average, 2 - 7%	0.31	0.34	0.36	0.4	0.43	0.47	0.50	0.56
Steep, >7%	0.35	0.39	0.41	0.45	0.48	0.52	0.54	0.58

length B exceeds $3T_c$, so that when T is 100, the proportion is approximately one third.

Validation of the runoff coefficients ($c_{(T)}$)

The purpose of validation is to test whether the model operates in the manner for which it was designed in "ways that were not explicitly built into the model" (Basson et al., 1994). Validation tests are necessary to convey confidence that the model works as expected. In order to validate the $c_{(T)}$ -values achieved in calibration, it was necessary to find some physical regional descriptor(s) on which to regress the coefficients. This was required so that the calibrated coefficients may be extended to un-gauged catchments.

Several regional descriptors were tested in combination with the $c_{(T)}$ -values to examine if a relationship existed on which to

TABLE 3
The mean and standard deviations of the ratio of the hydrograph time base-length B to the catchments' time of concentration T_c as a function of recurrence interval T . The proportion of B/T_c values above 3 in each group are given in the third row.

Recurrence interval T (years)	10	20	50	100	200
Mean of B/T_c ratios	1.92	2.06	2.25	2.40	2.56
Standard deviation	0.981	1.09	1.29	1.48	1.71
Proportion > 3	0.14	0.19	0.28	0.34	0.40

regress the coefficients. Descriptors such as catchment slope, mean annual precipitation (MAP), percentages of land coverage and Kovač's regional K -values (Kovač, 1988) were tested. From these analyses, no meaningful relationships between any of the descriptors tested and the c_{T_c} -coefficients were found. There were also no relationships found between parameters (multiplier and exponent) of a power-law function fitted to the c_{T_c} -values as a function of recurrence interval and regional descriptors. This result is in line with the comments of Pilgrim and Cordery (1993) for conditions in Australia, where the calibrated runoff coefficients did not show much sensitivity to catchment characteristics and indicate that the c -values are essentially functions of T and T_c as conjectured. Because there was no dependency observed between c -values and catchment properties, we were left with a problem: what values to use for validation?

It was decided to use the curves from Chow et al. (1988: 498), shown in Fig. 3, where it can be seen that the calibrated coefficients are generally lower than those of Chow et al. and the latter coefficients can be viewed as an approximate upper bounding set of curves. This choice, although conservative, was based on the premise that a practitioner will make a choice of the value of c based on catchment slope and land usage, knowing that it is bounded in the interval (0.1) and usually in the range 0.3 to 0.6.

Twenty-one catchments, which were not used in the calibration exercise and for which flood records were available, were selected for validation. These catchments ranged in size from 126 km² to 24 044 km². The flood records were modelled using a general extreme value (GEV) distribution in a previous study (Pegram and Parak, 2004) which was shown to be the most appropriate distribution generally speaking for flood peaks in the region. For these catchments, times of concentration (T_c) values were obtained from Petras and Du Plessis (1987) and representative design rainfall intensities from Smithers and Schulze (2002) in the same manner as for the calibration set. These data are summarised in Table A7 (Parts 1, 2 and 3) in the Appendix.

In order to obtain appropriate c -values from Chow et al. (1988: 498) for each catchment, it was necessary to relate the land coverage type and slope of each catchment with theirs (see Table 2 above). These catchment characteristics are given in Petras and Du Plessis (1987) where the percentages of land coverage for each catchment are catalogued as forest, dense bush wood, thin bush wood, cultivated land, grass and bare. At this stage it then became necessary to relate each catchment's coverage type (Petras and Du Plessis, 1987) to the generalised coverage types of Chow et al. (1988: 498). In order to easily accomplish this, several assumptions were made. They were:

- That the greatest percentage of land coverage (the modal type) was representative of the entire catchment
- That the following coverage types (from the descriptions of

Petras and Du Plessis (1987) and Chow et al. (1988) respectively) were equivalent (shown in Table 4 below).

TABLE 4
Equivalent land coverage types from the descriptions of Petras and Du Plessis (1987) and Chow et al. (1988: 498)

Equivalent land coverage types	
Actual catchment land coverage (as described in Petras and Du Plessis (1987))	c -coefficient land coverages (as listed in Chow et al. (1988: 498))
Forest	Forest/woodland
Dense bush wood	Forest/woodland
Thin bush wood	Forest/woodland
Cultivated land	Cultivated land
Grass	Pasture/range
Bare	Cultivated land

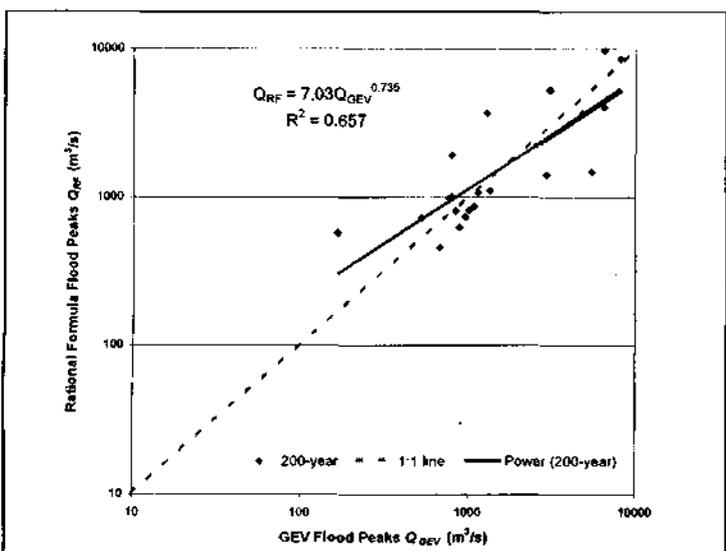
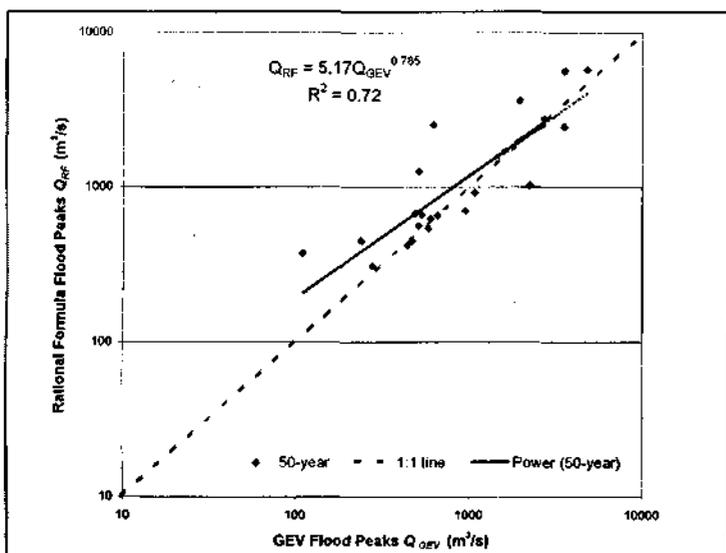
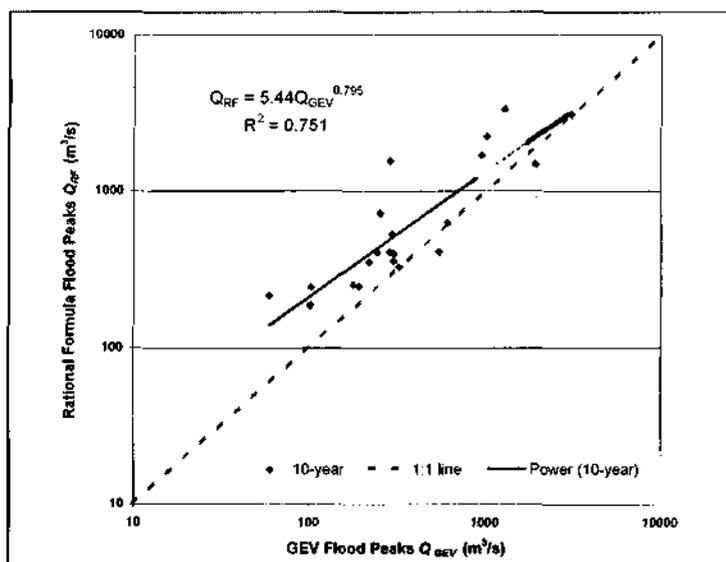
From the above procedure, design flood peaks were obtained using the rational formula method (Q_{RF}), i.e. a function of catchment area, design rainfall (of duration equal to the catchment time of concentration and the desired recurrence interval) and the runoff coefficients from Table 2. These design flood peaks were compared with the statistically modelled flood peaks (Q_{GEV}), from the same catchments, for the corresponding recurrence intervals. The results of this exercise, for the 10-, 50- and 200-year recurrence intervals are shown in Figs. 4, 5 and 6 respectively and are summarised for all recurrence intervals in Table 5.

Although there is a fairly large scatter around the trend-line in log-space in Figs. 4, 5 and 6, some conclusions can be drawn from this validation exercise.

TABLE 5
A summary of the power-law curves, of the form $Q_{RF} = aQ_{GEV}^b$, fitted to the graphs of Q_{RF} vs. Q_{GEV} (where Q_{RF} are the flood peaks obtained from the rational formula and Q_{GEV} are the statistically modelled flood peaks). The average ratio of Q_{RF}/Q_{GEV} for each recurrence interval is also given.

Recurrence interval T (years)	10	20	50	100	200
Factor: a	5.44	5.10	5.17	5.75	7.03
Exponent: b	0.795	0.798	0.785	0.766	0.735
R^2	0.751	0.746	0.726	0.699	0.657
Mean Q_{RF}/Q_{GEV}	1.84	1.64	1.42	1.31	1.21

It is evident from the 10-, 50- and 200-year validation graphs (shown in Figs. 4, 5 and 6 respectively) that the estimated rational formula flood peaks Q_{RF} tend to be larger than the GEV modelled flood peaks Q_{GEV} , especially for the lower magnitude floods, however, their trend-lines cross the 1:1 line at the larger flows – peaks at about 7 000 m³/s. This trend is also exhibited for the 20- and 100-year validation tests (the results of which are not shown here) and is confirmed in Table 5 where the average ratio of Q_{RF}/Q_{GEV} across all recurrence intervals is approximately 1.5, reducing from 1.84 for $T = 10$ to 1.21 for $T = 200$. This observation is to be expected since the c -values used to compute Q_{RF} from Chow et al. (1988: 498), were generally larger than the calibrated runoff coefficients obtained in this study (see Table 2 and Fig. 3). Although the R^2 -values are reasonable, the correlation is calculated in log-space and may disguise the fact that



some flow peak ratios are occasionally different by up to a factor of 5 (see Table A7, Part 3 in the Appendix for the full list of values). As a consequence, the c -values adopted for this validation exercise, from Chow et al., were treated as trial upper bound estimates, conceding that although consistent, the method is prone to error.

Discussion of results

Calibration

Calibration of the rational formula's runoff coefficients, using runhydrograph flood peak and volume pairs of given recurrence intervals, was performed with the intention of removing the subjectivity involved in this parameter's estimation in the design environment. Use was made of characteristic T -year flood peak and volume pairs together with T -year design rainfall intensities, as a function of the catchments time of concentration, in order to obtain the coefficients. The results of this exercise produced calibrated runoff coefficients, as a function of recurrence interval, which were scattered (see Fig. 3) around published values from Chow et al. (1988: 498). The calibrated values spread around the latter set of coefficients but were, in general, lower in magnitude (bar two catchments) and had gentler growths as a function of recurrence interval. Although this result did not produce a good match, the calibrated coefficients were sensible in magnitude. However, it was worrying to note that calibrated coefficients from six catchments (of the original 29) had a tendency to decrease in magnitude with increasing recurrence interval. This deviation from the norm is attributed to the fact that the flood runoff data (calculated using the runhydrograph method) had a gentler growth curve, as a function of recurrence interval, than the design rainfall data. It was found that the fitted c -values could not be regionalised in agree-

Figure 4 (top left)

A graph, for the purposes of validation, showing a plot in log space of the 10-year rational formula flood peaks Q_{RF} (using Chow et al.'s (1988) c -values as substitutes for calibrated runoff coefficients) vs. the 10-year GEV modelled flood peaks Q_{GEV} .

Figure 5 (middle left)

A graph, for the purposes of validation, showing a plot in log space of the 50-year rational formula flood peaks Q_{RF} (using Chow et al.'s (1988) c -values as substitutes for calibrated runoff coefficients) vs. the 50-year GEV modelled flood peaks Q_{GEV} .

Figure 6 (bottom left)

A graph, for the purposes of validation, showing a plot in log space of the 200-year rational formula flood peaks Q_{RF} (using Chow et al.'s (1988) c -values as substitutes for calibrated runoff coefficients) vs. the 200-year GEV modelled flood peaks Q_{GEV} .

ment with the conclusions of Pilgrim and Cordery (1993). Thus it is confirmed that c is a function of land-use, slope, T_c (through the design storm) and T . The fitted c -values (Fig. 3) were generally lower than those suggested by Chow et al. (1988: 498); it was therefore decided to accept the latter values for the purpose of validation, conscious of this discrepancy.

Hydrograph time base-length B

It was initially thought that this investigation would be able to produce entire design hydrographs (albeit in an idealised triangular form) from the rational formula since the flood data used (from the runhydrograph method) described characteristic peak and volume pairs for each catchment. It was hoped that the ratio of B to T_c (effectively a ratio of B to the time to peak of a rational formula hydrograph) would be consistent and that a particular outflow hydrograph could be prescribed with the use of this method. However, the results (see Table 3) indicate that, firstly, the average ratios are not constant across all recurrence intervals and, secondly, that the coefficients of variation are quite high (they range from 0.51 to 0.66). Also, the results shown in Table 3 exclude three catchments of area less than 130 km² as they gave ratios in excess of seven, however, several points are worth noting.

Firstly the base-lengths are, on average, 2.25 times the catchments' time of concentration across all recurrence intervals. This result is somewhat less than the length of the hydrograph suggested by Rooseboom et al. (1981) and Alexander (2002), which was $3T_c$, but Table 3 also indicates that a fair proportion of the calculated base lengths exceeded this number. As explained earlier, the hydrograph shape suggested by Rooseboom et al. (1981) was not meant to maintain continuity but was instead designed to be conservative. The hydrographs derived in this study are thus expected to have a smaller base-length as continuity is implicitly maintained, so the result is in line with expectation.

Secondly, the tendency of the base-length to increase with T is possibly due to the method employed by H&F in extracting their hydrographs and the non-linearity of the rainfall runoff process (abstractions reduce with T). As depicted in Fig. 7, H&F employed a truncation level for each catchment in order to extract independent hydrographs from their continuous records of streamflows. Flood volumes were obtained by extrapolating the rising limb and the recession limb of the discharge curves downwards towards zero flow from the first point below the truncation level which showed a reversal in slope. Depending on this level, a higher truncation level is likely to result in a reduction in the modelled volume when compared to the actual volume of the flood event. Thus it is likely that the base-lengths achieved in this study are smaller (as a function of T_c) for the smaller floods (more frequent events) than the base-lengths for the larger events, thus exhibiting the trend in Table 3.

Finally, it is interesting to examine the relationship between B and T_c using a linear rainfall-runoff model as a comment on the values appearing in Table 3. If a constant (pulsed) input of rainfall of intensity i (in mm/h) on a catchment of area A (in km²) lasts for the time of concentration T_c (hours), the total volume of rain that falls is $V = 1000 \cdot i \cdot T_c \cdot A$ (in m³). The average rate of flow onto the catchment is $1000 \cdot i \cdot A$ (in m³/h) and the peak outflow Q must be a fraction of this, say $\alpha \cdot 1000 \cdot i \cdot A$ (m³/h), where $0 < \alpha < 1$ (α is a factor related to the closeness of the peak to its asymptotic value as defined by its nearness to equilibrium). The base-length of the equivalent triangular hydrograph is thus $B = V/Q = 2 \cdot T_c / \alpha$ (in hours). If there are no losses, the maximum peak that occurs at T_c can only be approaching equilibrium asymptotically,

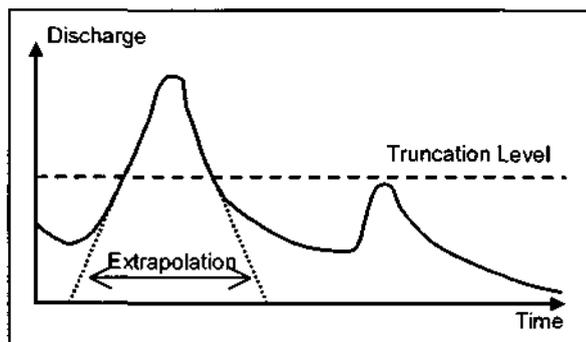


Figure 7
The method employed by Hiemstra and Francis (1979) to extract independent hydrographs from a continuous flow record, showing that a lower truncation level is likely to provide a bigger volume.

so α has to be chosen close to 1. If $\alpha = 0.9$, then it turns out that $B \approx 2.2T_c$, which is close to the average ratio (determined from Table 3 above).

Validation

The validation exercise was necessary to test whether the calibrated coefficients behaved in the probabilistic manner for which they were designed, i.e. to predict design floods of magnitudes equivalent to those derived from a statistical analysis of flood records from that site. However, since it was shown that c is not dependent on physical properties nor location of the un-gauged catchments, c -values from Chow et al. (1988: 498), which are a function of T_c , catchment slope and land-use characteristics, were substituted for the calibrated coefficients as approximate upper bound values. Based on this substitution, the validation exercise was ultimately reduced to a test of whether the c -values from Chow et al. (or possibly some other summary values) could provide reasonable design flood estimates such as those obtained from a statistical distribution (such as the GEV) fit to historical flood data.

The result of this exercise showed that the floods estimated using the substitute c -values from Chow et al. (1988) produced floods from the rational formula that were, on average, approximately 1.5 times larger than the floods estimated from the statistical distributions of the historical data (see Table 5 and Figs. 4, 5 and 6), with a tendency to overestimate for lower flood peaks and T . This result is in line with expectation as the substitute c -values from Chow et al. (1988) were adopted as upper bound estimates. Given that, in order to make use of the coefficients of Chow et al., a crude matching of land coverage types was performed (see Table 4), this result is relatively pleasing especially since the catchments used in validation ranged in size from small to large (170 to 24 000 km² – see Table A7). The precision of the method is of course still low, as indicated by the spread of results in Figs. 4, 5 and 6, and relies heavily on the judgement of the practitioner.

Conclusion

The rational formula, which is possibly the simplest rainfall-runoff flood estimation technique available was reviewed by means of calibrating the most uncertain variable of the formula, i.e. the runoff coefficient c . The "data set" used to achieve this was the set of runhydrographs produced by Hiemstra and

Francis (1979). The results of the calibration were reasonably encouraging, producing c -coefficients that were scattered around, but generally lower than, those offered by Chow et al. (1988: 498), whose precision is not known. It was discovered that the fitted $c_{(A,Tc)}$ -coefficients of this investigation did not show any variation with catchment characteristics, in line with Australian experience (Pilgrim and Cordery, 1993), and hence validation of these values at other sites was only possible using land-use and average slope of validation catchments together with recurrence interval as guides for the choice of c -values. It was thus decided to use the c -values from Chow et al. as approximate upper bound estimates of the fitted $c_{(A,Tc)}$ -coefficients in validation. In order to use their values, a match of land coverage types was required. The results of the validation were as expected, producing floods from the rational formula that were on average 1.5 times larger than the floods estimated from a statistical analysis of the validation set (not used for calibration), but with a wide scatter. Of minor importance, it was discovered that the time base-lengths of the derived triangular hydrographs of this investigation were approximately between 1.9 and 2.6 times the catchment's time of concentration, depending on the recurrence interval of the flood, lower than suggested elsewhere. It can be concluded, from the results of this investigation, that the rational formula is a simple, consistent, approximate tool when used in its probabilistic frame-work and although not suitable as a stand-alone design tool for flood estimation, can be useful as a quick check method for calculating flood hydrographs for large catchments as it is for small.

Acknowledgements

The first author would like to acknowledge the post-graduate funding provided by the South African Department of Transport, which made this research towards improved design flood estimation techniques possible.

References

- ADAMSON PT (1981) Southern African Storm Rainfall. Tech. Rep. TR102, Department of Water Affairs, Pretoria.
- ALEXANDER WJR (1990) *Flood Hydrology for Southern Africa*. SANCOLD, Pretoria.
- ALEXANDER WJR (2002) The Standard Design Flood – Theory and Practice. Report, Dept of Civil Engineering, University of Pretoria, Pretoria.
- BASSON MS, ALLEN RB, PEGRAM GGS and VAN ROOYEN JA (1994) *Probabilistic Management of Water resource and Hydro-power Systems*. Water Resources Publications, Colorado.
- CHOW VT, MAIDMENT DR and MAYS LR (1988) *Applied Hydrology*. McGraw-Hill, New York.
- DE MICHELE D, SALVADORI G, CANOSSO M, PETACCIA A and ROSSO R (2005) Bivariate statistical approach to check adequacy of dam spillway. *J. Hydrol. Eng. ASCE* 10 (1) 50-57
- FRANCIS DM (1979) The Runhydrograph Applied. M.Sc. Thesis. Dept of Civil Engineering, University of Natal, Durban.
- FSR (1975) *Flood Studies Report. Vol. 2 Meteorological Studies*. Natural Environment Research Council, London.
- GORGENS AHM (2002) Design Flood Hydrology: Design and Rehabilitation of Dams. G Basson (ed.) Institute for Water and Environmental Engineering, Department of Civil Engineering, University of Stellenbosch, South Africa. 460-524.
- HIEMSTRA LAV (1972) Runhydrographs – A new technique in hydrograph generation. *Proc. Int. Symp. on Modelling Techniques in Water Resources Systems*. Ottawa, Canada. 1 205-214.
- HIEMSTRA LAV (1973) Runhydrographs for the sizing of dam spillways and the minimum reservoir capacities. *Proc. International Commission on Large Dams*. Madrid, Spain, 1 217-237.
- HIEMSTRA LAV (1974) Runhydrographs from Poisson-generated run lengths. *J. Hydraul. Div. ASCE* 100 1617-1630.
- HIEMSTRA LAV and FRANCIS DM (1979) The Runhydrograph – Theory and Application for Flood Predictions. Water Research Commission, Pretoria.
- HIEMSTRA LAV, ZUCCHINI WS and PEGRAM GGS (1976) A method of finding the family of runhydrographs for given return periods. *J. Hydrol.* 30 95-103.
- HRU REPORT NO 1/72 (1972) Design Flood Determination in South Africa. DC Midgley (ed.) Hydrological Research Unit, Department of Civil Engineering, University of Witwatersrand, South Africa.
- INSTITUTE OF ENGINEERS AUSTRALIA (1987) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. DH Pilgrim (ed.) Canberra, Australia.
- KIRPICH ZP (1940) Time of concentration of small agricultural watersheds. *Civ. Eng.* 10 (6) 362.
- KOVACS Z (1988) *Regional Maximum Flood Peaks in Southern Africa*. Tech. Rep. TR 137, Department of Water Affairs, Pretoria.
- PEGRAM GGS (2003) Rainfall, rational formula and regional maximum flood – some scaling links. Keynote Paper. *Aust. J. Water Resour.* 7 (1) 29-39.
- PEGRAM GGS and PARAK M (2004) A review of the regional maximum flood and rational formula using geomorphological information and observed floods. *Water SA* 30 (3) 377-392.
- PETRAS V and DU PLESSIS PH (1987) Catalogue of Hydrological Parameters. Flood Studies. Technical Note No. 6, Department of Water Affairs, Pretoria.
- PILGRIM DH and CORDERY I (1993) Chapter 9: Flood Runoff In: DR Maidment (ed.) *Handbook of Hydrology*. McGraw-Hill, New York.
- ROOSEBOOM A, BASSON MS, LOOTS CH, WIGGET JH and BOSMAN J (1981) *National Transport Commission Road Drainage Manual (1st edn.)* National Transport Commission, Directorate Land Transport, Pretoria
- SMITHERS JC and SCHULZE RE (2002) Design Rainfall and Flood Estimation in South Africa. WRC Report 1060/1/03, Water Research Commission, South Africa.

No.	Station	River	Lat. (deg & min)	Long. (deg & min)	Area (km ²)	Length of record (years)	Quality of data	No. of hydro- graphs	Trun- cation level (m ³ /s)	Deg. of trunc. %	ln (peaks)		ln (vol)		Cross correl- ation coeff.	Chi- square bi- variate	Accep- tance level Chi- square
											Mean	St. Dev.	Mean	St. dev.			
1	A2M01	Krokodil	25 44	27 52	2 907	18	A	20	323.4	43.51	5.964	1.133	7.812	1.727	0.843	13.36	70.1
2	A2M02	Magalies	25 44	27 51	1 206	18	A	17	116.6	64.36	4.282	1.297	5.896	1.461	0.907	18.24	70.1
3	A2M03	Hex	25 46	27 17	494	19	A	48	79.8	12.90	5.061	0.601	6.001	0.839	0.846	8.64	88.0
4	A2M12	Krokodil	25 49	27 55	2 586	36	D	37	109.6	49.87	4.699	0.698	6.432	1.065	0.750	12.97	88.0
5	A3M01	Klein Marico	25 32	26 06	1 002	33	A	33	52.4	73.44	3.338	0.992	5.307	0.776	0.704	20.31	88.0
6	B2M01	Bronkhorstspuit	25 48	28 46	1 585	47	A	42	145.9	36.14	5.242	0.731	7.664	1.062	0.923	21.70	88.0
7	B4M03	Steelpoort	25 01	29 32	2 271	10	A	9	76.4	45.57	4.423	0.787	7.871	0.555	0.748	8.59	54.4
8	B7M04	Klaserie	24 33	31 02	130	10	D	24	41.9	33.07	4.048	0.716	5.848	1.163	0.586	7.98	70.1
9	C1M01	Vaal	26 57	29 16	8 254	54	A	56	433.8	44.73	6.162	0.678	10.158	0.907	0.847	14.39	88.0
10	C3M04	Drv Hartz	27 34	24 43	8 039	24	A	18	47.1	25.81	4.241	0.600	6.561	1.187	0.676	21.39	70.1
11	C4M01	Groot Vet	28 29	26 40	5 504	24	B	40	325.3	40.22	5.962	0.718	9.359	0.663	0.827	11.69	88.0
12	C4M02	Vet	27 51	25 54	17 550	14	A	41	217.5	13.79	5.899	0.473	10.392	0.645	0.839	11.27	88.0
13	C5M03	Modder	29 10	26 35	1 650	36	A	35	247.4	9.04	6.189	0.504	8.806	0.778	0.889	6.86	88.0
14	C5M04	Modder	28 51	26 11	5 012	28	A	24	449.6	12.44	6.863	0.652	9.640	0.838	0.879	6.49	70.1
15	C5M10	Kromellenboog	29 50	25 38	1 994	17	A	28	43.4	11.61	4.686	0.765	7.353	1.055	0.893	9.96	70.1
16	C5M12	Riet	29 39	25 59	2 383	21	D	20	127.7	11.46	5.547	0.577	8.491	0.758	0.877	5.07	70.1
17	C5M15	Modder	28 48	26 06	6 545	18	A	18	298.9	18.30	6.252	0.611	9.278	0.894	0.892	6.07	70.1
18	C7M01	Renoster	27 16	27 11	5 255	25	A	14	218.8	38.02	5.671	0.931	9.566	0.924	0.892	145.16	54.4
19	C9M03	Vaal	28 31	24 42	10 8081	16	C	19	382.6	9.01	6.715	0.569	11.599	1.112	0.892	5.62	70.1
20	C9M06	Vaal	27 39	25 35	10 2384	22	A	23	519.6	35.57	6.554	0.815	11.565	1.151	0.754	9.18	70.1
21	D1M05	Oranje	30 02	28 30	10 891	29	A	26	574	59.87	6.195	0.631	10.593	0.599	0.683	8.04	70.1
22	D2M05	Caledon	28 53	27 54	3 815	14	A	37	171.5	26.04	5.538	0.613	9.724	0.731	0.692	7.56	88.0
23	D3M05	Oranje	29 48	24 26	91 994	19	A	22	2 302.6	64.51	7.509	0.625	12.75	0.476	0.439	12.85	70.1
24	D5M01	Renoster	31 39	20 37	2 129	26	E	31	82.19	6.86	5.174	0.510	7.918	0.769	0.839	9.81	70.1
25	D5M04	Sak	31 39	21 46	5 799	31	E	36	149.56	14.05	5.511	0.466	8.421	0.712	0.651	6.69	88.0
26	D6M02	Brak	30 07	23 34	6 360	16	A	36	73.9	16.47	4.925	0.638	7.688	0.995	0.760	9.77	88.0
27	E2M02	Doring	32 30	19 32	5 778	37	C	40	210.9	46.64	5.414	0.741	9.310	0.826	0.935	32.01	88.0
28	G1M02	Vier-en-twintig	33 08	19 04	186	10	A	9	261.4	9.12	6.010	0.331	7.388	0.683	0.577	4.92	54.4
29	H1M06	Bree	33 25	19 16	754	10	A	12	360.8	32.31	6.083	0.425	8.962	0.384	0.573	10.83	54.4
30	H1M07	Wit	33 34	19 09	83	10	A	27	171.5	66.75	4.918	0.524	6.884	0.847	0.608	18.19	70.1
31	H7M04	Huis	33 55	20 43	26	10	A	9	10.2	24.98	2.881	0.837	5.118	1.223	0.707	5.34	54.4
32	J2M03	Gamka	33 32	21 39	17 941	19	A	25	81.7	20.37	5.120	0.867	7.718	1.667	0.765	10.47	70.1
33	J3M04	Olifants	33 29	23 02	4 330	44	E	29	172.7	19.70	5.723	0.671	7.802	0.731	0.852	10.57	70.1
34	L7M02	Groot Vet	33 20	24 21	25 587	19	A	35	165.6	10.65	5.703	0.523	8.560	0.935	0.716	8.10	88.0
35	Q1M01	Groot Vis	31 54	25 29	9 150	43	A	30	242.1	60.84	5.273	0.787	7.399	1.314	0.882	89.34	70.1
36	Q7M01	Groot Vis	32 57	25 49	18 954	23	A	72	243.3	18.11	6.059	0.620	9.133	0.698	0.808	18.83	107.8
37	Q7M02	Groot Vis	32 43	25 51	18 436	26	A	19	275.6	16.35	6.230	0.623	9.221	0.880	0.858	6.58	70.1
38	Q9M10	Groot Vis	33 13	26 52	29 376	26	A	13	671.4	22.34	7.192	0.898	10.513	0.942	0.899	7.25	54.4
39	Q9M12	Groot Vis	33 06	26 27	23 041	24	B	35	86.6	22.88	5.242	1.052	8.615	1.315	0.916	11.21	88.0
40	T3M02	Kinira	30 29	28 37	2 100	12	A	15	207.3	8.01	5.671	0.238	8.953	0.929	0.494	7.01	54.4
42	W4A03	Pongola	27 25	31 31	5 843	16	A	13	655.7	36.17	6.710	0.636	9.328	0.700	0.877	40.60	54.4
42	W5M05	Hlelo	26 50	30 44	751	19	A	25	41.6	49.55	3.739	1.003	5.831	1.541	0.666	14.84	70.1
43	X2A02	Wit	25 19	31 03	176	20	B	20	48.3	10.51	4.511	0.503	6.194	1.231	0.736	6.85	70.1

No.	Station	Lat. (dec deg)	Long (dec deg)	Area (km ²)	Time of conc. (h)	Runhydrograph design flood (used for calibration)			Design rainfall (used in rational formula for calibration)				Calibrated c -coefficient, $c_{(T)}$	Comment
						Peak (m ³ /s)	Vol x 10 ⁶ (m ³)	Hydro. base-length (h)	a	b	c	Intensity i mm/h		
1	A2M03	25.77	27.28	494	6.4	510.8	7.5	8.2	52.2	0.225	0.776	12.37	0.301	
2	A2M12	25.82	27.92	2 586	18	331.4	12.1	20.2	46.8	0.237	0.763	5.16	0.089	
3	A3M01	25.53	26.10	1 002	8.7	204.3	3.4	9.3	48.2	0.210	0.790	8.73	0.084	
4	B2M01	25.80	28.77	1 585	18.1	531.2	34.4	36.0	53.9	0.227	0.773	5.74	0.210	
5	B4M03	25.02	29.53	2 271	19.6	249.7	20.5	45.5	41.7	0.240	0.760	4.34	0.091	
6	B7M04	24.55	31.03	130	3.7	215.9	10.8	27.7	69.1	0.240	0.760	25.56	0.234	Anomalous
7	C1M01	26.95	29.27	8 254	74	1 404.4	396.7	156.9	42.8	0.229	0.772	1.55	0.396	
8	C4M01	28.48	26.67	5 504	34	1 524.0	147.6	53.8	45.5	0.200	0.800	2.71	0.368	
9	C4M02	27.85	25.90	17 550	111	968.5	444.5	255.0	44.0	0.218	0.782	1.11	0.179	Anomalous
10	C5M03	29.17	26.58	1 650	18.3	843.5	56.0	36.9	41.5	0.227	0.773	4.39	0.419	
11	C5M04	28.85	26.18	5 012	38	1 807.6	125.4	38.5	41.1	0.226	0.774	2.46	0.528	
12	C5M12	29.65	25.98	2 383	23	479.1	39.8	46.2	38.3	0.220	0.780	3.31	0.218	
13	C5M15	28.80	26.10	6 545	43	1 106.8	116.7	58.6	40.1	0.225	0.776	2.17	0.280	
14	C7M01	27.27	27.18	5 255	57	666.7	117.2	97.7	45.1	0.221	0.779	1.93	0.236	
15	D1M05	30.03	28.50	10 891	60	1 339.0	372.4	154.5	44.1	0.204	0.796	1.69	0.261	
16	D5M01	31.65	20.62	2 129	27	331.0	25.5	42.8	32.6	0.172	0.828	2.13	0.263	Anomalous
17	D5M04	31.65	21.77	5 799	28	418.4	36.5	48.4	31.0	0.177	0.823	2.00	0.130	Anomalous
18	E2M02	32.50	19.53	5 778	30	835.1	172.0	114.4	20.7	0.194	0.806	1.34	0.389	
19	H1M06	33.42	19.27	754	7.6	765.5	46.5	33.7	29.8	0.355	0.645	8.05	0.454	
20	H1M07	33.57	19.15	83	2.4	475.9	26.4	30.8	42.9	0.399	0.601	25.37	0.814	Anomalous
21	H7M04	33.92	20.72	26	2.3	43.0	2.2	28.1	38.3	0.301	0.699	21.42	0.278	
22	J2M03	33.53	21.65	17 941	42	567.6	84.7	82.9	31.6	0.186	0.814	1.51	0.076	
23	J3M04	33.48	23.03	4 330	23	500.1	15.0	16.7	30.3	0.211	0.789	2.55	0.163	
24	Q1M01	31.90	25.48	9 150	18	683.0	47.7	38.8	29.9	0.209	0.791	3.04	0.089	
25	Q9M10	33.22	26.87	29 376	108	1 962.9	199.4	56.4	38.1	0.289	0.711	1.37	0.176	
26	Q9M12	33.10	26.45	23 041	85	1 114.7	182.4	90.9	33.5	0.307	0.693	1.54	0.113	
27	T3M02	30.48	28.62	2 100	26	370.0	71.7	107.7	45.1	0.207	0.793	3.40	0.186	Anomalous
28	W4A03	27.42	31.52	5 843	31	1 833.4	98.1	29.7	54.0	0.259	0.741	4.23	0.267	
29	W5M05	26.83	30.73	751	17.8	230.8	16.8	40.4	57.6	0.229	0.771	6.25	0.177	

TABLE A3
Data used for the calibration of the rational formula's c-coefficient for the 20-year case

No.	Station	Lat. (dec deg)	Long (dec deg)	Area (km ²)	Time of conc. (h)	Runhydrograph design flood (used for calibration)			Design rainfall (used in rational formula for calibration)				Cali- brated c- coeff- icient, c _(T)	Comment
						Peak (m ³ /s)	Vol x 10 ⁶ (m ³)	Hydro. base- length (h)	a	b	c	Inten- sity i mm/h		
1	A2M03	25.77	27.28	494	6.4	597.9	9.3	8.7	60.7	0.225	0.776	14.39	0.303	
2	A2M12	25.82	27.92	2 586	18	404.2	16.3	22.4	55.2	0.237	0.763	6.08	0.093	
3	A3M01	25.53	26.10	1 002	8.7	258.6	4.1	8.8	55.6	0.210	0.790	10.06	0.092	
4	B2M01	25.80	28.77	1 585	18.1	680.3	49.3	40.2	63.7	0.227	0.773	6.79	0.228	
5	B4M03	25.02	29.53	2 271	19.6	317.7	24.2	42.4	47.3	0.240	0.760	4.93	0.102	
6	B7M04	24.55	31.03	130	3.7	255.2	14.1	30.7	82.2	0.240	0.760	30.39	0.233	Anomalous
7	C1M01	26.95	29.27	8 254	74	1 719.7	520.1	168.0	49.6	0.229	0.772	1.79	0.419	
8	C4M01	28.48	26.67	5 504	34	1 841.8	175.8	53.0	52.3	0.200	0.800	3.12	0.386	
9	C4M02	27.85	25.90	17 550	111	1 090.5	522.7	266.3	50.9	0.218	0.782	1.28	0.175	Anomalous
10	C5M03	29.17	26.58	1 650	18.3	1 019.7	75.1	40.9	47.8	0.227	0.773	5.06	0.440	
11	C5M04	28.85	26.18	5 012	38	2 334.5	174.2	41.5	47.3	0.226	0.774	2.83	0.592	
12	C5M12	29.65	25.98	2 383	23	594.7	52.9	49.5	44.1	0.220	0.780	3.82	0.235	
13	C5M15	28.80	26.10	6 545	43	1 374.5	160.1	64.7	46.2	0.224	0.776	2.50	0.302	
14	C7M01	27.27	27.18	5 255	57	976.4	171.2	97.4	52.0	0.221	0.779	2.23	0.300	
15	D1M05	30.03	28.50	10 891	60	1 590.5	438.5	153.2	51.4	0.204	0.796	1.98	0.266	
16	D5M01	31.65	20.62	2 129	27	394.6	33.2	46.8	38.7	0.172	0.828	2.52	0.264	Anomalous
17	D5M04	31.65	21.77	5 799	28	486.5	45.9	52.5	36.6	0.177	0.823	2.36	0.128	Anomalous
18	E2M02	32.50	19.53	5 778	30	1 041.5	220.0	117.4	24.0	0.194	0.806	1.55	0.420	
19	H1M06	33.42	19.27	754	7.6	864.7	51.9	33.3	33.4	0.355	0.645	9.03	0.457	
20	H1M07	33.57	19.15	83	2.4	528.0	31.2	32.8	48.4	0.399	0.601	28.62	0.800	Anomalous
21	H7M04	33.92	20.72	26	2.3	57.4	3.3	32.1	46.4	0.301	0.699	25.92	0.307	
22	J2M03	33.53	21.65	17 941	42	740.8	141.4	106.0	37.9	0.186	0.814	1.81	0.082	
23	J3M04	33.48	23.03	4 330	23	667.0	20.6	17.1	36.6	0.211	0.789	3.09	0.180	
24	Q1M01	31.90	25.48	9 150	18	869.3	71.4	45.6	34.6	0.209	0.791	3.51	0.097	
25	Q9M10	33.22	26.87	29 376	108	3 080.4	320.0	57.7	46.3	0.289	0.711	1.66	0.227	
26	Q9M12	33.10	26.45	23 041	85	1 533.4	271.7	98.4	39.2	0.307	0.693	1.81	0.133	
27	T3M02	30.48	28.62	2 100	26	398.5	95.8	133.5	52.6	0.207	0.793	3.97	0.172	Anomalous
28	W4A03	27.42	31.52	5 843	31	2 287.1	125.1	30.4	64.7	0.259	0.742	5.07	0.278	
29	W5M05	26.83	30.73	751	17.8	299.3	25.0	46.4	68.4	0.229	0.771	7.43	0.193	

TABLE A4
Data used for the calibration of the rational formula's c-coefficient for the 50-year case

No.	Station	Lat. (dec deg)	Long (dec deg)	Area (km ²)	Time of conc. (h)	Runhydrograph design flood (used for calibration)			Design rainfall (used in rational formula for calibration)				Calibrated c-coefficient, c_m	Comment
						Peak (m ³ /s)	Vol x 10 ⁶ (m ³)	Hydro. base-length (h)	a	b	c	Intensity /mm/h		
1	A2M03	25.77	27.28	494	6.4	718.4	12.1	9.3	72.6	0.224	0.776	17.21	0.304	
2	A2M12	25.82	27.92	2 586	18	507.2	23.1	25.3	67.2	0.237	0.763	7.41	0.095	
3	A3M01	25.53	26.10	1 002	8.7	341.6	5.1	8.3	65.3	0.210	0.790	11.82	0.104	
4	B2M01	25.80	28.77	1 585	18.1	895.0	73.4	45.6	78.2	0.227	0.773	8.33	0.244	
5	B4M03	25.02	29.53	2 271	19.6	416.6	29.3	39.1	56.4	0.240	0.760	5.88	0.112	
6	B7M04	24.55	31.03	130	3.7	310.7	19.4	34.8	101.3	0.294	0.706	40.21	0.214	Anomalous
7	C1M01	26.95	29.27	8 254	74	2 165.0	707.8	181.6	58.9	0.228	0.772	2.13	0.444	
8	C4M01	28.48	26.67	5 504	34	2 296.0	215.5	52.1	61.7	0.200	0.800	3.68	0.409	
9	C4M02	27.85	25.90	17 550	111	1 253.8	632.2	280.1	60.2	0.218	0.782	1.51	0.170	Anomalous
10	C5M03	29.17	26.58	1 650	18.3	1 252.0	103.1	45.7	56.3	0.227	0.773	5.96	0.458	
11	C5M04	28.85	26.18	5 012	38	3 069.0	247.6	44.8	55.7	0.226	0.774	3.34	0.660	
12	C5M12	29.65	25.98	2 383	23	751.4	72.0	53.2	51.9	0.220	0.780	4.50	0.252	
13	C5M15	28.80	26.10	6 545	43	1 743.1	226.7	72.3	54.5	0.224	0.776	2.95	0.325	
14	C7M01	27.27	27.18	5 255	57	1 462.1	255.6	97.1	61.6	0.221	0.779	2.64	0.379	
15	D1M05	30.03	28.50	10 891	60	1 938.1	529.0	151.6	61.7	0.204	0.796	2.37	0.270	
16	D5M01	31.65	20.62	2 129	27	478.8	44.5	51.6	46.9	0.172	0.828	3.06	0.264	Anomalous
17	D5M04	31.65	21.77	5 799	28	574.9	59.3	57.3	44.2	0.177	0.823	2.84	0.125	Anomalous
18	E2M02	32.50	19.53	5 778	30	1 338.5	291.0	120.8	28.2	0.194	0.806	1.82	0.459	
19	H1M06	33.42	19.27	754	7.6	992.4	58.8	32.9	38.0	0.355	0.645	10.27	0.461	
20	H1M07	33.57	19.15	83	2.4	598.6	38.2	35.5	55.7	0.399	0.601	32.88	0.790	Anomalous
21	H7M04	33.92	20.72	26	2.3	78.7	5.3	37.2	58.1	0.301	0.699	32.48	0.356	
22	J2M03	33.53	21.65	17 941	42	999.6	251.6	139.8	46.9	0.186	0.814	2.24	0.090	
23	J3M04	33.48	23.03	4 330	23	900.4	28.5	17.6	45.9	0.211	0.789	3.87	0.194	
24	Q1M01	31.90	25.48	9 150	18	1 142.1	112.6	54.8	40.9	0.209	0.791	4.15	0.108	
25	Q9M10	33.22	26.87	29 376	108	4 835.3	513.5	59.0	58.6	0.289	0.711	2.10	0.282	
26	Q9M12	33.10	26.45	23 041	85	2 201.4	427.0	107.8	47.2	0.307	0.693	2.17	0.158	
27	T3M02	30.48	28.62	2 100	26	432.4	131.8	169.3	63.0	0.207	0.793	4.76	0.156	Anomalous
28	W4A03	27.42	31.52	5 843	31	2 917.6	163.6	31.1	80.8	0.259	0.741	6.33	0.284	
29	W5M05	26.83	30.73	751	17.8	404.7	39.8	54.6	84.3	0.229	0.771	9.15	0.212	

TABLE A5
Data used for the calibration of the rational formula's c-coefficient for the 100-year case

No.	Station	Lat. (dec deg)	Long (dec deg)	Area (km ²)	Time of conc. (h)	Runhydrograph design flood (used for calibration)			Design rainfall (used in rational formula for calibration)				Calibrated c-coefficient, c_m	Comment
						Peak (m ³ /s)	Vol x 10 ⁶ (m ³)	Hydro. base-length (h)	a	b	c	Intensity i mm/h		
1	A2M03	25.77	27.28	494	6.4	815.3	14.4	9.8	82.2	0.224	0.776	19.48	0.305	
2	A2M12	25.82	27.92	2 586	18	591.8	29.2	27.4	77.3	0.237	0.763	8.51	0.097	
3	A3M01	25.53	26.10	1 002	8.7	414.4	6.0	8.0	72.8	0.210	0.790	13.19	0.113	
4	B2M01	25.80	28.77	1 585	18.1	1 074.7	95.8	49.5	90.3	0.227	0.773	9.62	0.254	
5	B4M03	25.02	29.53	2 271	19.6	500.0	33.4	37.1	60.7	0.240	0.760	6.33	0.125	
6	B7M04	24.55	31.03	130	3.7	356.2	24.3	37.9	117.5	0.240	0.760	43.45	0.227	Anomalous
7	C1M01	26.95	29.27	8 254	74	2 530.6	872.1	191.5	66.4	0.228	0.772	2.40	0.460	
8	C4M01	28.48	26.67	5 504	34	2 672.3	247.9	51.5	68.9	0.200	0.800	4.11	0.425	
9	C4M02	27.85	25.90	17 550	111	1 381.1	721.3	290.1	67.5	0.218	0.782	1.70	0.167	Anomalous
10	C5M03	29.17	26.58	1 650	18.3	1 432.5	126.9	49.2	63.0	0.227	0.773	6.66	0.469	
11	C5M04	28.85	26.18	5 012	38	3 668.3	311.4	47.2	62.3	0.226	0.774	3.73	0.706	
12	C5M12	29.65	25.98	2 383	23	876.1	88.1	55.8	58.0	0.220	0.780	5.02	0.264	
13	C5M15	28.80	26.10	6 545	43	2 040.3	285.4	77.7	60.9	0.224	0.776	3.29	0.341	
14	C7M01	27.27	27.18	5 255	57	1 899.5	331.4	96.9	69.2	0.221	0.779	2.97	0.438	
15	D1M05	30.03	28.50	10 891	60	2 218.0	601.2	150.6	70.0	0.204	0.796	2.69	0.272	
16	D5M01	31.65	20.62	2 129	27	544.4	54.0	55.1	53.4	0.172	0.828	3.49	0.264	Anomalous
17	D5M04	31.65	21.77	5 799	28	642.4	70.2	60.8	50.2	0.177	0.823	3.23	0.123	Anomalous
18	E2M02	32.50	19.53	5 778	30	1 586.0	351.6	123.2	31.4	0.194	0.806	2.03	0.487	
19	H1M06	33.42	19.27	754	7.6	1 089.1	63.9	32.6	41.4	0.355	0.645	11.20	0.464	
20	H1M07	33.57	19.15	83	2.4	654.2	44.1	37.5	61.0	0.399	0.601	36.07	0.787	Anomalous
21	H7M04	33.92	20.72	26	2.3	97.0	7.1	40.9	68.1	0.301	0.699	38.05	0.353	
22	J2M03	33.53	21.65	17 941	42	1 223.2	370.9	168.4	54.4	0.186	0.814	2.59	0.095	
23	J3M04	33.48	23.03	4 330	23	1 091.9	35.2	17.9	53.8	0.211	0.789	4.53	0.200	
24	Q1M01	31.90	25.48	9 150	18	1 373.6	153.2	62.0	45.8	0.209	0.791	4.66	0.116	
25	Q9M10	33.22	26.87	29 376	108	6 417.6	691.0	59.8	69.0	0.289	0.711	2.48	0.318	
26	Q9M12	33.10	26.45	23 041	85	2 811.4	579.7	114.5	53.5	0.307	0.693	2.47	0.178	
27	T3M02	30.48	28.62	2 100	26	456.6	163.0	198.3	71.5	0.207	0.793	5.40	0.145	Anomalous
28	W4A03	27.42	31.52	5 843	31	3 429.6	195.4	31.7	94.6	0.259	0.741	7.42	0.285	
29	W5M05	26.83	30.73	751	17.8	497.8	54.6	61.0	97.8	0.229	0.771	10.62	0.225	

TABLE A6
Data used for the calibration of the rational formula's c-coefficient for the 200-year case

No.	Station	Lat. (dec deg)	Long (dec deg)	Area (km ²)	Time of conc. (h)	Runhydrograph design flood (used for calibration)			Design rainfall (used in rational formula for calibration)				Calibrated c-coefficient, c _m	Comment
						Peak (m ³ /s)	Vol x 10 ⁶ (m ³)	Hydro. base-length (h)	a	b	c	Intensity i mm/h		
1	A2M03	25.77	27.28	494	6.4	918.2	17.0	10.3	92.4	0.224	0.776	21.89	0.306	
2	A2M12	25.82	27.92	2 586	18	683.2	36.4	29.6	88.2	0.237	0.763	9.72	0.098	
3	A3M01	25.53	26.10	1 002	8.7	497.3	6.9	7.7	80.5	0.210	0.790	14.58	0.123	
4	B2M01	25.80	28.77	1 585	18.1	1 271.5	122.3	53.4	103.5	0.227	0.773	11.03	0.262	
5	B4M03	25.02	29.55	2 271	19.6	591.9	37.6	35.3	66.7	0.240	0.760	6.95	0.135	
6	B7M04	24.55	31.03	130	3.7	405.2	29.9	41.0	135.4	0.240	0.760	50.07	0.224	Anomalous
7	C1M01	26.95	29.27	8 254	74	2 925.1	1 058.5	201.0	74.3	0.228	0.772	2.68	0.476	
8	C4M01	28.48	26.67	5 504	34	3 080.9	282.7	51.0	76.4	0.200	0.800	4.55	0.442	
9	C4M02	27.85	25.90	17 550	111	1 512.7	816.6	299.9	75.2	0.218	0.782	1.89	0.164	Anomalous
10	C5M03	29.17	26.58	1 650	18.3	1 619.5	153.4	52.6	69.8	0.227	0.773	7.38	0.479	
11	C5M04	28.85	26.18	5 012	38	4 312.8	383.5	49.4	69.0	0.226	0.774	4.14	0.749	
12	C5M12	29.65	25.98	2 383	23	1 007.6	105.8	58.3	64.2	0.220	0.780	5.56	0.274	
13	C5M15	28.80	26.10	6 545	43	2 356.6	352.5	83.1	67.5	0.225	0.776	3.65	0.355	
14	C7M01	27.27	27.18	5 255	57	2 406.6	419.2	96.8	77.2	0.221	0.779	3.31	0.498	
15	D1M05	30.03	28.50	10 891	60	2 515.3	677.5	149.6	79.0	0.204	0.796	3.04	0.274	
16	D5M01	31.65	20.62	2 129	27	612.4	64.5	58.5	60.2	0.172	0.828	3.93	0.264	Anomalous
17	D5M04	31.65	21.77	5 799	28	711.2	82.1	64.1	56.5	0.177	0.823	3.64	0.121	Anomalous
18	E2M02	32.50	19.53	5 778	30	1 856.2	418.9	125.4	34.7	0.194	0.806	2.24	0.516	
19	H1M06	33.42	19.27	754	7.6	1 187.0	69.1	32.3	44.8	0.355	0.645	12.12	0.468	
20	H1M07	33.57	19.15	83	2.4	712.1	50.6	39.5	66.5	0.399	0.601	39.28	0.786	Anomalous
21	H7M04	33.92	20.72	26	2.3	117.3	9.4	44.7	79.1	0.301	0.699	44.17	0.368	
22	J2M03	33.53	21.65	17 941	42	1 474.0	530.9	200.1	62.4	0.186	0.814	2.98	0.099	
23	J3M04	33.48	23.03	4 330	23	1 298.7	42.5	18.2	62.4	0.211	0.789	5.26	0.205	
24	Q1M01	31.90	25.48	9 150	18	1 629.7	203.8	69.5	50.9	0.209	0.791	5.17	0.124	
25	Q9M10	33.22	26.87	29 376	108	8 251.8	899.5	60.6	80.8	0.289	0.711	2.90	0.349	
26	Q9M12	33.10	26.45	23 041	85	3 526.9	769.6	121.2	60.3	0.307	0.693	2.78	0.198	
27	T3M02	30.48	28.62	2 100	26	480.0	198.0	229.2	80.7	0.207	0.793	6.09	0.135	Anomalous
28	W4A03	27.42	31.52	5 843	31	3 977.5	230.1	32.1	110.2	0.259	0.741	8.64	0.284	
29	W5M05	26.83	30.73	751	17.8	604.2	73.6	67.7	112.7	0.228	0.771	12.23	0.237	

TABLE A7 (Part 1)
Data used for validation

Num.	Station	River	Lat. (dec deg)	Long. (dec deg)	Area (km ²)	Time of Conc. (hours)	Percentage land coverage						Runoff coefficients <i>c</i> from Chow et al. (1988)				
							For- est	Bush Wood		Cult. Land	Grass	Bare	10- year	20- year	50- year	100- year	200- year
								Dense	Thin								
1	A4M02	Mokolo	24.28	28.09	1 777	18.1	0	25	39	26	7	3	0.28	0.31	0.35	0.39	0.42
2	A6M06	Klein-nyl	24.70	28.41	168	4.4	0	28	34	29	6	3	0.28	0.31	0.35	0.39	0.42
3	C3M03	Harts	27.58	24.75	10990	78	0	0	28	59	4	9	0.36	0.39	0.43	0.47	0.51
4	C5M08	Riet	29.81	26.21	593	11.9	0	0	7	25	22	46	0.36	0.39	0.43	0.47	0.51
5	C8M01	Wilge	27.27	28.32	15673	122	0	6	8	50	24	12	0.36	0.39	0.43	0.47	0.51
6	C8M03	Cornelis	27.84	28.96	806	19.2	0	3	4	42	34	17	0.36	0.39	0.43	0.47	0.51
7	D1M01	Stormbergspruit	31.00	26.34	2 397	19.9	0	0	46	13	14	27	0.28	0.31	0.35	0.39	0.42
8	D1M04	Stormbergspruit	31.40	26.37	348	9.1	0	0	50	30	7	13	0.28	0.31	0.35	0.39	0.42
9	D2M01	Caledon	29.72	26.98	13421	106	0	4	7	73	11	5	0.36	0.39	0.43	0.47	0.51
10	E2M03	Doring	31.90	18.69	24044	59	0	0	16	16	0	68	0.36	0.39	0.43	0.47	0.51
11	G1M08	Klein-berg	33.31	19.08	395	4	5	20	5	48	15	7	0.36	0.39	0.43	0.47	0.51
12	H3M01	Kingna	33.79	20.13	611	9.5	0	0	25	15	0	60	0.36	0.39	0.43	0.47	0.51
13	Q7M03	Groot-vis	32.78	25.84	18534	59	0	0	30	8	0	62	0.36	0.39	0.43	0.47	0.51
14	Q9M04	Kat	32.56	26.69	404	6.3	4	8	24	16	32	16	0.38	0.41	0.45	0.49	0.52
15	Q9M08	Kat	32.71	26.59	748	12.7	7	3	25	19	31	15	0.38	0.41	0.45	0.49	0.52
16	R1M01	Tyume	32.76	26.86	238	6.2	10	10	15	20	30	15	0.38	0.41	0.45	0.49	0.52
17	T3M04	Mzimhlava	30.57	29.43	1029	18.8	5	5	14	46	20	10	0.36	0.39	0.43	0.47	0.51
18	V2M02	Mooi	29.22	29.99	937	18.9	20	10	0	25	43	2	0.30	0.33	0.37	0.41	0.46
19	V6M02	Tugela	28.75	30.44	12862	48	1	24	0	30	42	3	0.30	0.33	0.37	0.41	0.46
20	W5M06	Swartwater	27.11	30.83	180	5	15	13	0	7	64	1	0.30	0.33	0.37	0.41	0.46
21	X2M10	Noordkaap	25.61	30.88	126	3.3	45	15	0	10	28	2	0.28	0.31	0.35	0.39	0.42

TABLE A7 (Part 2)
Design rainfall data used for validation

No.	Station	b	c	a					Rainfall Intensity, $i=ad^c$ (mm/hr)				
				10-year	20-year	50-year	100-year	200-year	10-year	20-year	50-year	100-year	200-year
1	A4M02	0.218	0.783	50.19	58.40	69.88	79.13	88.86	5.21	6.06	7.25	8.21	9.22
2	A6M06	0.217	0.783	49.87	58.07	69.45	78.64	88.37	16.37	19.05	22.79	25.80	28.98
3	C3M03	0.188	0.812	48.66	56.27	66.31	73.96	81.78	1.41	1.63	1.92	2.15	2.37
4	C5M08	0.220	0.780	38.17	44.28	52.68	59.38	66.52	5.53	6.42	7.64	8.61	9.64
5	C8M01	0.234	0.766	42.93	49.42	58.35	65.43	72.77	1.08	1.25	1.47	1.65	1.84
6	C8M03	0.231	0.769	42.23	48.28	56.45	62.94	69.62	4.35	4.97	5.81	6.48	7.17
7	D1M01	0.214	0.786	35.20	40.01	46.24	50.85	55.48	3.36	3.82	4.41	4.85	5.29
8	D1M04	0.240	0.760	36.85	42.07	48.99	54.31	59.66	6.88	7.86	9.15	10.14	11.15
9	D2M01	0.226	0.774	41.02	47.57	56.61	63.85	71.49	1.11	1.29	1.53	1.73	1.94
10	E2M03	0.201	0.799	24.06	27.80	32.73	36.52	40.36	0.93	1.07	1.26	1.40	1.55
11	G1M08	0.327	0.673	26.19	29.38	33.45	36.49	39.44	10.31	11.56	13.16	14.35	15.52
12	H3M01	0.278	0.722	29.63	36.07	45.70	53.97	63.29	5.83	7.09	8.98	10.61	12.44
13	Q7M03	0.295	0.705	32.07	37.50	45.10	51.20	57.66	1.81	2.12	2.54	2.89	3.25
14	Q9M04	0.252	0.748	37.78	44.20	53.17	60.33	67.99	9.53	11.15	13.41	15.22	17.15
15	Q9M08	0.285	0.715	30.99	36.22	43.57	49.48	55.71	5.03	5.89	7.08	8.04	9.05
16	R1M01	0.258	0.742	38.84	45.44	54.62	62.00	69.88	10.03	11.74	14.11	16.02	18.05
17	T3M04	0.221	0.779	51.08	60.15	73.32	84.35	96.45	5.19	6.12	7.46	8.58	9.81
18	V2M02	0.223	0.777	51.33	59.81	71.72	81.45	91.87	5.22	6.09	7.30	8.29	9.35
19	V6M02	0.228	0.772	57.56	69.04	86.15	100.91	117.52	2.90	3.47	4.33	5.08	5.91
20	W5M06	0.227	0.773	57.15	67.96	83.76	97.13	111.94	16.48	19.59	24.15	28.01	32.28
21	X2M10	0.225	0.776	62.92	74.80	92.18	106.94	123.24	24.93	29.63	36.53	42.37	48.83

TABLE A7 (Part 3)
Flood data used for validation

No.	Station	Rational formula flood peaks, Q_{RF} (m ³ /s), using design rainfall and					GEV modelled flood peaks of recorded events Q_{GEV} (m ³ /s)				
		c-values from Chow et al. (1988)					10-year	20-year	50-year	100-year	200-year
		10-year	20-year	50-year	100-year	200-year					
1	A4M02	719.5	913.9	1 252.1	1 579.5	1 930.1	255.0	359.0	516.0	652.0	806.0
2	A6M06	213.9	271.7	372.2	469.6	573.8	59.0	79.0	110.0	137.0	168.0
3	C3M03	1 552.6	1962.2	2 524.9	3 079.5	3 678.0	291.0	397.0	622.0	893.0	1 301.0
4	C5M08	328.2	416.3	540.9	666.6	806.4	325.0	432.0	585.0	712.0	851.0
5	C8M01	1 699.9	2 141.5	2 759.3	3 383.6	4 067.3	959.0	1 515.0	2 719.0	4 197.0	6 456.0
6	C8M03	350.4	438.2	559.7	682.0	814.6	220.0	321.0	514.0	727.0	1 024.0
7	D1M01	626.0	776.6	1 027.5	1 258.9	1 493.9	616.0	1 144.0	2 246.0	3 557.0	5 505.0
8	D1M04	186.3	232.2	309.6	382.3	457.1	101.0	156.0	279.0	437.0	687.0
9	D2M01	1 490.9	1 890.9	2 457.4	3 029.3	3 663.7	1 939.0	2 653.0	3 543.0	4 185.0	4 804.0
10	E2M03	2 227.1	2 813.1	3 616.8	4 407.9	5 261.7	1 029.0	1 389.0	1 956.0	2 470.0	3 073.0
11	G1M08	407.2	499.5	620.9	740.3	864.6	288.0	400.0	603.0	813.0	1 092.0
12	H3M01	356.0	474.0	655.5	846.6	1 072.2	304.0	436.0	661.0	881.0	1 156.0
13	Q7M03	3 351.5	4 290.2	5 630.4	6 986.4	8 505.1	1 289.0	2 022.0	3 561.0	5 399.0	8 141.0
14	Q9M04	406.3	517.2	677.0	836.8	1 007.9	245.0	340.0	493.0	635.0	805.0
15	Q9M08	397.4	505.6	661.8	818.4	985.3	305.0	401.0	539.0	652.0	775.0
16	R1M01	252.0	320.7	419.7	518.8	624.9	178.0	269.0	442.0	632.0	895.0
17	T3M04	534.5	688.8	916.5	1 152.4	1 423.8	297.0	534.0	1 083.0	1 801.0	2 963.0
18	V2M02	408.0	529.3	703.2	884.6	1 108.5	548.0	716.0	954.0	1 149.0	1 357.0
19	V6M02	3 104.3	4 144.6	5 728.4	7 435.4	9 618.5	3 096.0	3 791.0	4 790.0	5 620.0	6 523.0
20	W5M06	247.2	327.3	446.8	574.2	735.1	192.0	292.0	474.0	662.0	970.0
21	X2M10	244.3	317.0	447.4	578.3	725.0	102.0	146.0	240.0	355.0	530.0

APPENDIX C

C-1. Code to resolve a D8 flow direction raster into a D4 flow direction raster in a clockwise direction:

```
% Code to resolve the flow direction into a "D4" raster (clockwise) %

% Values of data contained in ASCII array
ncols = 60;
nrows = 114;
nodata_value = -9999;

% Loading the data file, input file
D8_Data = load('C:/ARCGIS/Text_Files/D8.txt');

% Create a matrix of nodata_values which will be replaced with D4 flow
% direction values
for a=1:nrows
    for b=1:ncols
        D4_Flow_Dir(a,b)=nodata_value;
    end;
end;

% Loop through the data set
for i=1:nrows
    for j=1:ncols

        Flow_Direc = D8_Data(i,j);

        if D8_Data(i,j)==nodata_value
            continue
        end;

        if Flow_Direc==128
            D4_Flow_Dir(i,j)=1;
        elseif Flow_Direc==2
            D4_Flow_Dir(i,j)=4;
        elseif Flow_Direc==8
            D4_Flow_Dir(i,j)=16;
        elseif Flow_Direc==32
            D4_Flow_Dir(i,j)=64;
        else
            D4_Flow_Dir(i,j)=Flow_Direc;
        end;

    end;
end;

% Output file directory
dlmwrite('C:/ARCGIS/Text_Files/D4_Flow_Dir_1.txt', D4_Flow_Dir, '\t')
```

C-2. Code to resolve a D8 flow direction raster into a D4 flow direction raster in an anti-clockwise direction:

```
% Code to resolve the flow direction into a "D4" raster (anti-
% clockwise)

% Values of data contained in ASCII array
ncols = 60;
nrows = 114;
nodata_value = -9999;

% Loading the data file, input file
D8_Data = load('C:/ARCGIS/Text_Files/D8.txt');

% Create a matrix of nodata_values which will be replaced with D4 flow
% direction values
for a=1:nrows
    for b=1:ncols
        D4_Flow_Dir(a,b)=nodata_value;
    end;
end;

% Loop through the data set
for i=1:nrows
    for j=1:ncols

        Flow_Direc = D8_Data(i,j);

        if D8_Data(i,j)==nodata_value
            continue
        end;

        if Flow_Direc==128
            D4_Flow_Dir(i,j)=64;
        elseif Flow_Direc==2
            D4_Flow_Dir(i,j)=1;
        elseif Flow_Direc==8
            D4_Flow_Dir(i,j)=4;
        elseif Flow_Direc==32
            D4_Flow_Dir(i,j)=16;
        else
            D4_Flow_Dir(i,j)=Flow_Direc;
        end;

    end;
end;

% Output file directory
dlmwrite('C:/ARCGIS/Text_Files/D4_Flow_Dir_1.txt', D4_Flow_Dir,'\t')
```

C-3. Code to resolve a D8 flow direction raster into a D4 flow direction raster in the direction of the steepest neighbour:

```

% Code to resolve the flow direction into a "D4" raster (steepest
% neighbor)

% Values of data contained in ASCII array
ncols = 64;
nrows = 123;
nodata_value = -9999;
problem_value = -100000;

% Loading the data file, input file
D8_Data = load('C:/ARCGIS/Code/Text_Files/D8_2.txt');
Height_Data = load('C:/ARCGIS/Code/Text_Files/fill_1km_dem_2.txt');

% Create a matrix of nodata_values which will be replaced with D4 flow
% direction values
for a=1:nrows
    for b=1:ncols
        D4_Flow_Dir(a,b)=nodata_value;
    end;
end;

% Loop through the data set
for i=2:(nrows-1)
    for j=2:(ncols-1)

        Height = Height_Data(i,j);
        Height_Up = Height_Data(i-1,j);
        Height_Down = Height_Data(i+1,j);
        Height_Right = Height_Data(i,j+1);
        Height_Left = Height_Data(i,j-1);

        Flow_Direc = D8_Data(i,j);

        if D8_Data(i,j)==nodata_value
            continue
        end;

        if Flow_Direc==128
            if Height_Up<Height_Right && Height_Up~=nodata_value
                D4_Flow_Dir(i,j)=64;
            elseif Height_Right~=nodata_value
                D4_Flow_Dir(i,j)=1;
            else
                D4_Flow_Dir(i,j)=-100000;
            end;
        end;

        if Flow_Direc==2
            if Height_Right<Height_Down && Height_Right~=nodata_value
                D4_Flow_Dir(i,j)=1;
            elseif Height_Down~=nodata_value
                D4_Flow_Dir(i,j)=4;
            else
                D4_Flow_Dir(i,j)=-100000;
            end;
        end;
    end;
end;

```

```
if Flow_Direc==8
    if Height_Down<Height_Left && Height_Down~=nodata_value
        D4_Flow_Dir(i,j)=4;
    elseif Height_Left~=nodata_value
        D4_Flow_Dir(i,j)=16;
    else
        D4_Flow_Dir(i,j)=-100000;
    end;
end;

if Flow_Direc==32
    if Height_Left<Height_Up && Height_Left~=nodata_value
        D4_Flow_Dir(i,j)=16;
    elseif Height_Up~=nodata_value
        D4_Flow_Dir(i,j)=64;
    else
        D4_Flow_Dir(i,j)=-100000;
    end;
end;

if Flow_Dir==64 || Flow_Dir==1 || Flow_Dir~=
end;

end;

% Output file directory
dlmwrite('C:/ARCGIS/Code/Text_Files/D4_Flow_Dir_1.txt',
D4_Flow_Dir, '\t')
```

C-4. Code to find the erroneous flow direction codes from the resolved "D4" raster

```

% Code to find the erroneous flow direction codes from the resolved
% "D4" raster

% Values of data contained in ASCII array
ncols = 64;
nrows = 123;
nodata_value = -9999;
problem_value = -100000;

% Loading the data file, input file
D4_Data = load('C:/ARCGIS/Code/Text_Files/D4_Flow_Dir_2.txt');

% Create a matrix of nodata_values which will flag the erroneous
% values
for a=1:nrows
    for b=1:ncols
        Flag(a,b)=nodata_value;
    end;
end;

% Loop through the data set
for i=2:(nrows-1)
    for j=2:(ncols-1)

        Dir = D4_Data(i,j);
        Dir_Up = D4_Data(i-1,j);
        Dir_Down = D4_Data(i+1,j);
        Dir_Right = D4_Data(i,j+1);
        Dir_Left = D4_Data(i,j-1);

        if Dir==64 && Dir_Up==4
            Flag(i,j)=10;
            Flag(i-1,j)=10;
        end;

        if Dir==4 && Dir_Down==64
            Flag(i,j)=10;
            Flag(i+1,j)=10;
        end;

        if Dir==1 && Dir_Right==16
            Flag(i,j)=20;
            Flag(i,j+1)=20;
        end;

        if Dir==16 && Dir_Left==1
            Flag(i,j)=20;
            Flag(i,j-1)=20;
        end;

        if Dir==problem_value
            Flag(i,j)=30;
        end;

    end;
end;

```

```
end;
```

```
% Output file directory
```

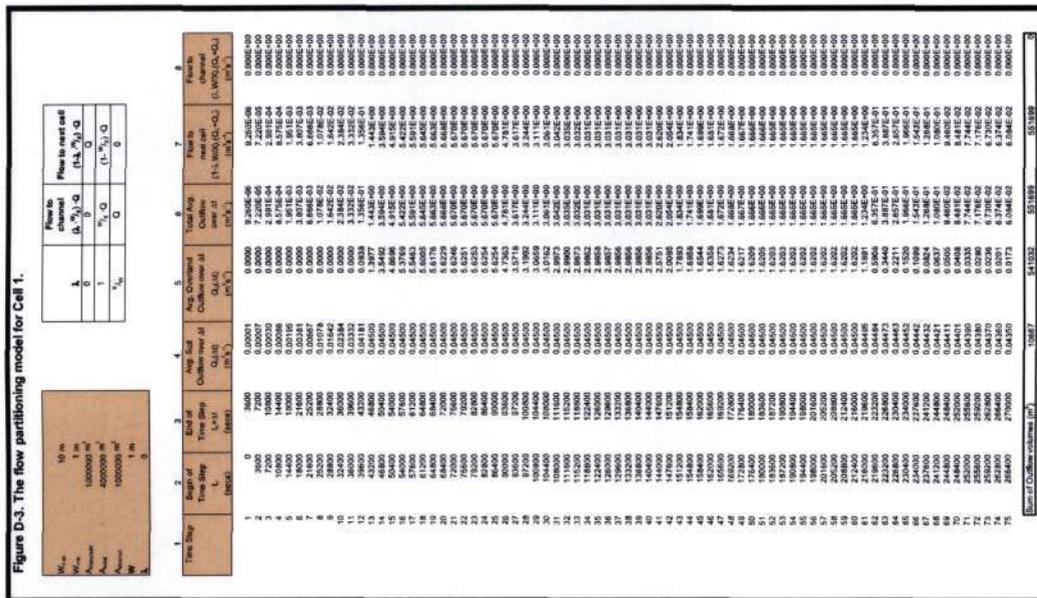
```
dlmwrite('C:/ARCGIS/Code/Text_Files/Flag.txt', Flag, '\t')
```

APPENDIX D

Appendix D presents a printout of the test catchment of Chapter 7 as it is simulated with the TOPKAPI model on Microsoft Excel™.

Figure D-2. The overlaid store model for Cell 1.

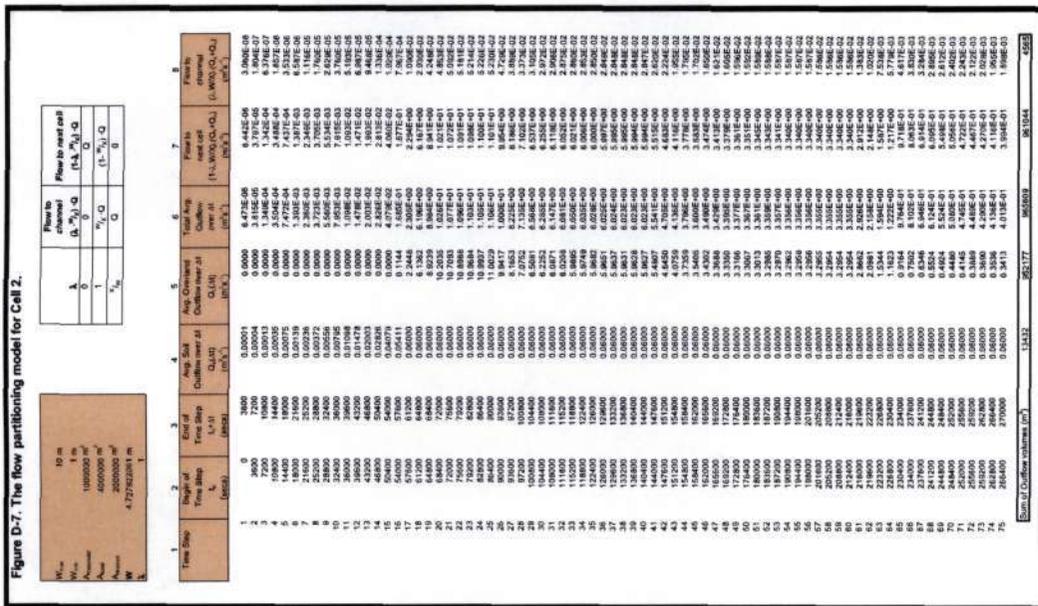
Time Step	Begin of Time Step (hr)	End of Time Step (hr)	Saturation Excess (mm)	Overland Flow Rate (mm/h)	Max. Store Depth (mm)	Outflow @ 1-hr (mm/h)	Outflow @ 1-hr (mm/h)	Time Step (hr)	End of Time Step (hr)	Max. Store Depth (mm)	Outflow @ 1-hr (mm/h)	Outflow @ 1-hr (mm/h)	Time Step (hr)	End of Time Step (hr)	Max. Store Depth (mm)	Outflow @ 1-hr (mm/h)	Outflow @ 1-hr (mm/h)
1	0	3600	0	0.0000	0.0000	0.0000	0.0000	0	3600	0.0000	0.0000	0.0000	0	3600	0.0000	0.0000	0.0000
2	3600	7200	0	0.0000	0.0000	0.0000	0.0000	0	7200	0.0000	0.0000	0.0000	0	7200	0.0000	0.0000	0.0000
3	7200	10800	0	0.0000	0.0000	0.0000	0.0000	0	10800	0.0000	0.0000	0.0000	0	10800	0.0000	0.0000	0.0000
4	10800	14400	0	0.0000	0.0000	0.0000	0.0000	0	14400	0.0000	0.0000	0.0000	0	14400	0.0000	0.0000	0.0000
5	14400	18000	0	0.0000	0.0000	0.0000	0.0000	0	18000	0.0000	0.0000	0.0000	0	18000	0.0000	0.0000	0.0000
6	18000	21600	0	0.0000	0.0000	0.0000	0.0000	0	21600	0.0000	0.0000	0.0000	0	21600	0.0000	0.0000	0.0000
7	21600	25200	0	0.0000	0.0000	0.0000	0.0000	0	25200	0.0000	0.0000	0.0000	0	25200	0.0000	0.0000	0.0000
8	25200	28800	0	0.0000	0.0000	0.0000	0.0000	0	28800	0.0000	0.0000	0.0000	0	28800	0.0000	0.0000	0.0000
9	28800	32400	0	0.0000	0.0000	0.0000	0.0000	0	32400	0.0000	0.0000	0.0000	0	32400	0.0000	0.0000	0.0000
10	32400	36000	0	0.0000	0.0000	0.0000	0.0000	0	36000	0.0000	0.0000	0.0000	0	36000	0.0000	0.0000	0.0000
11	36000	39600	0	0.0000	0.0000	0.0000	0.0000	0	39600	0.0000	0.0000	0.0000	0	39600	0.0000	0.0000	0.0000
12	39600	43200	0	0.0000	0.0000	0.0000	0.0000	0	43200	0.0000	0.0000	0.0000	0	43200	0.0000	0.0000	0.0000
13	43200	46800	0	0.0000	0.0000	0.0000	0.0000	0	46800	0.0000	0.0000	0.0000	0	46800	0.0000	0.0000	0.0000
14	46800	50400	0	0.0000	0.0000	0.0000	0.0000	0	50400	0.0000	0.0000	0.0000	0	50400	0.0000	0.0000	0.0000
15	50400	54000	0	0.0000	0.0000	0.0000	0.0000	0	54000	0.0000	0.0000	0.0000	0	54000	0.0000	0.0000	0.0000
16	54000	57600	0	0.0000	0.0000	0.0000	0.0000	0	57600	0.0000	0.0000	0.0000	0	57600	0.0000	0.0000	0.0000
17	57600	61200	0	0.0000	0.0000	0.0000	0.0000	0	61200	0.0000	0.0000	0.0000	0	61200	0.0000	0.0000	0.0000
18	61200	64800	0	0.0000	0.0000	0.0000	0.0000	0	64800	0.0000	0.0000	0.0000	0	64800	0.0000	0.0000	0.0000
19	64800	68400	0	0.0000	0.0000	0.0000	0.0000	0	68400	0.0000	0.0000	0.0000	0	68400	0.0000	0.0000	0.0000
20	68400	72000	0	0.0000	0.0000	0.0000	0.0000	0	72000	0.0000	0.0000	0.0000	0	72000	0.0000	0.0000	0.0000
21	72000	75600	0	0.0000	0.0000	0.0000	0.0000	0	75600	0.0000	0.0000	0.0000	0	75600	0.0000	0.0000	0.0000
22	75600	79200	0	0.0000	0.0000	0.0000	0.0000	0	79200	0.0000	0.0000	0.0000	0	79200	0.0000	0.0000	0.0000
23	79200	82800	0	0.0000	0.0000	0.0000	0.0000	0	82800	0.0000	0.0000	0.0000	0	82800	0.0000	0.0000	0.0000
24	82800	86400	0	0.0000	0.0000	0.0000	0.0000	0	86400	0.0000	0.0000	0.0000	0	86400	0.0000	0.0000	0.0000
25	86400	90000	0	0.0000	0.0000	0.0000	0.0000	0	90000	0.0000	0.0000	0.0000	0	90000	0.0000	0.0000	0.0000
26	90000	93600	0	0.0000	0.0000	0.0000	0.0000	0	93600	0.0000	0.0000	0.0000	0	93600	0.0000	0.0000	0.0000
27	93600	97200	0	0.0000	0.0000	0.0000	0.0000	0	97200	0.0000	0.0000	0.0000	0	97200	0.0000	0.0000	0.0000
28	97200	100800	0	0.0000	0.0000	0.0000	0.0000	0	100800	0.0000	0.0000	0.0000	0	100800	0.0000	0.0000	0.0000
29	100800	104400	0	0.0000	0.0000	0.0000	0.0000	0	104400	0.0000	0.0000	0.0000	0	104400	0.0000	0.0000	0.0000
30	104400	108000	0	0.0000	0.0000	0.0000	0.0000	0	108000	0.0000	0.0000	0.0000	0	108000	0.0000	0.0000	0.0000
31	108000	111600	0	0.0000	0.0000	0.0000	0.0000	0	111600	0.0000	0.0000	0.0000	0	111600	0.0000	0.0000	0.0000
32	111600	115200	0	0.0000	0.0000	0.0000	0.0000	0	115200	0.0000	0.0000	0.0000	0	115200	0.0000	0.0000	0.0000
33	115200	118800	0	0.0000	0.0000	0.0000	0.0000	0	118800	0.0000	0.0000	0.0000	0	118800	0.0000	0.0000	0.0000
34	118800	122400	0	0.0000	0.0000	0.0000	0.0000	0	122400	0.0000	0.0000	0.0000	0	122400	0.0000	0.0000	0.0000
35	122400	126000	0	0.0000	0.0000	0.0000	0.0000	0	126000	0.0000	0.0000	0.0000	0	126000	0.0000	0.0000	0.0000
36	126000	129600	0	0.0000	0.0000	0.0000	0.0000	0	129600	0.0000	0.0000	0.0000	0	129600	0.0000	0.0000	0.0000
37	129600	133200	0	0.0000	0.0000	0.0000	0.0000	0	133200	0.0000	0.0000	0.0000	0	133200	0.0000	0.0000	0.0000
38	133200	136800	0	0.0000	0.0000	0.0000	0.0000	0	136800	0.0000	0.0000	0.0000	0	136800	0.0000	0.0000	0.0000
39	136800	140400	0	0.0000	0.0000	0.0000	0.0000	0	140400	0.0000	0.0000	0.0000	0	140400	0.0000	0.0000	0.0000
40	140400	144000	0	0.0000	0.0000	0.0000	0.0000	0	144000	0.0000	0.0000	0.0000	0	144000	0.0000	0.0000	0.0000
41	144000	147600	0	0.0000	0.0000	0.0000	0.0000	0	147600	0.0000	0.0000	0.0000	0	147600	0.0000	0.0000	0.0000
42	147600	151200	0	0.0000	0.0000	0.0000	0.0000	0	151200	0.0000	0.0000	0.0000	0	151200	0.0000	0.0000	0.0000
43	151200	154800	0	0.0000	0.0000	0.0000	0.0000	0	154800	0.0000	0.0000	0.0000	0	154800	0.0000	0.0000	0.0000
44	154800	158400	0	0.0000	0.0000	0.0000	0.0000	0	158400	0.0000	0.0000	0.0000	0	158400	0.0000	0.0000	0.0000
45	158400	162000	0	0.0000	0.0000	0.0000	0.0000	0	162000	0.0000	0.0000	0.0000	0	162000	0.0000	0.0000	0.0000
46	162000	165600	0	0.0000	0.0000	0.0000	0.0000	0	165600	0.0000	0.0000	0.0000	0	165600	0.0000	0.0000	0.0000
47	165600	169200	0	0.0000	0.0000	0.0000	0.0000	0	169200	0.0000	0.0000	0.0000	0	169200	0.0000	0.0000	0.0000
48	169200	172800	0	0.0000	0.0000	0.0000	0.0000	0	172800	0.0000	0.0000	0.0000	0	172800	0.0000	0.0000	0.0000
49	172800	176400	0	0.0000	0.0000	0.0000	0.0000	0	176400	0.0000	0.0000	0.0000	0	176400	0.0000	0.0000	0.0000
50	176400	180000	0	0.0000	0.0000	0.0000	0.0000	0	180000	0.0000	0.0000	0.0000	0	180000	0.0000	0.0000	0.0000
51	180000	183600	0	0.0000	0.0000	0.0000	0.0000	0	183600	0.0000	0.0000	0.0000	0	183600	0.0000	0.0000	0.0000
52	183600	187200	0	0.0000	0.0000	0.0000	0.0000	0	187200	0.0000	0.0000	0.0000	0	187200	0.0000	0.0000	0.0000
53	187200	190800	0	0.0000	0.0000	0.0000	0.0000	0	190800	0.0000	0.0000	0.0000	0	190800	0.0000	0.0000	0.0000
54	190800	194400	0	0.0000	0.0000	0.0000	0.0000	0	194400	0.0000	0.0000	0.0000	0	194400	0.0000	0.0000	0.0000
55	194400	198000	0	0.0000	0.0000	0.0000	0.0000	0	198000	0.0000	0.0000	0.0000	0	198000	0.0000	0.0000	0.0000
56	198000	201600	0	0.0000	0.0000	0.0000	0.0000	0	201600	0.0000	0.0000	0.0000	0	201600	0.0000	0.0000	0.0000
57	201600	205200	0	0.0000	0.0000	0.0000	0.0000	0	205200	0.0000	0.0000	0.0000	0	205200	0.0000	0.0000	0.0000
58	205200	208800	0	0.0000	0.0000	0.0000	0.0000	0	208800	0.0000	0.0000	0.0000	0	208800	0.0000	0.0000	0.0000
59	208800	212400	0	0.0000	0.0000	0.0000	0.0000	0	212400	0.0000	0.0000	0.0000	0	212400	0.0000	0.0000	0.0000
60	212400	216000	0	0.0000	0.0000	0.0000	0.0000	0	2								

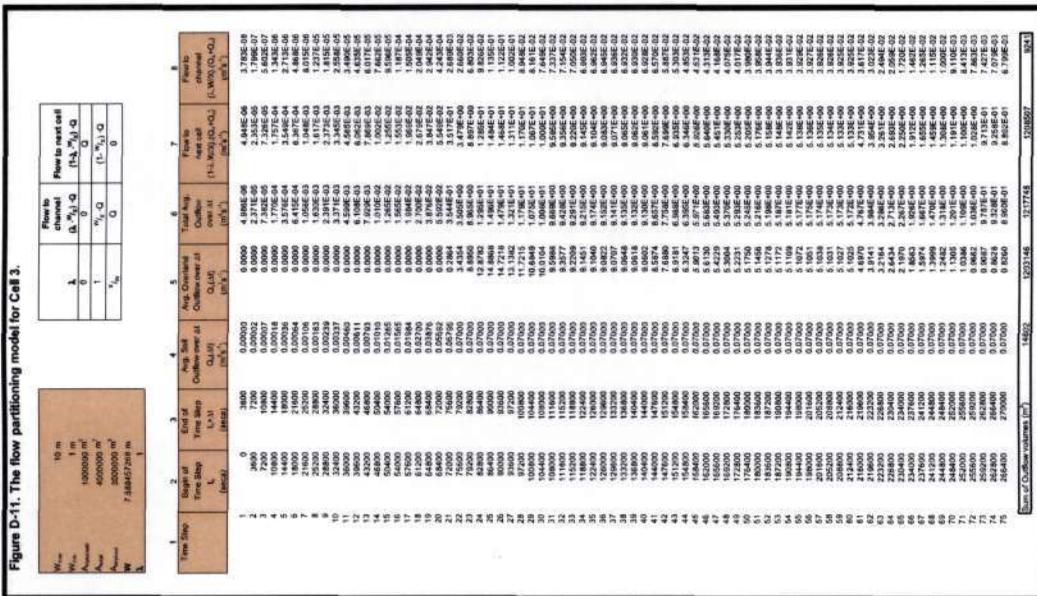


Continuity sum = 1.61844274

Check analytical solution to the convergence differential equation for the channel points

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Time Step (hr)	Step of Time Step (hr)	Wet Front Cell (m)	Upstream Channel (m)	Total Peak Channel (m)	Maximum Stored @ h (m)	Outflow @ h (m/s)	Y (m)	β	α	β^2/α^2	A	B	C	ϕ_h	ϕ_v	Time Step Interval (hr)	Storage @ h (m ³)	Maximum Storage @ h (m ³)	Outflow @ h (m/s)	Reg. Outflow at end of cell (m/s)
1	0	3600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	7200	7200	0	0.00000
2	0	7200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	14400	14400	0	0.00000
3	0	10800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	21600	21600	0	0.00000
4	0	14400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	28800	28800	0	0.00000
5	0	18000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	36000	36000	0	0.00000
6	0	21600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	43200	43200	0	0.00000
7	0	25200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	50400	50400	0	0.00000
8	0	28800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	57600	57600	0	0.00000
9	0	32400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	64800	64800	0	0.00000
10	0	36000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	72000	72000	0	0.00000
11	0	39600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	79200	79200	0	0.00000
12	0	43200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	86400	86400	0	0.00000
13	0	46800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	93600	93600	0	0.00000
14	0	50400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	100800	100800	0	0.00000
15	0	54000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	108000	108000	0	0.00000
16	0	57600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	115200	115200	0	0.00000
17	0	61200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	122400	122400	0	0.00000
18	0	64800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	129600	129600	0	0.00000
19	0	68400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	136800	136800	0	0.00000
20	0	72000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	144000	144000	0	0.00000
21	0	75600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	151200	151200	0	0.00000
22	0	79200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	158400	158400	0	0.00000
23	0	82800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	165600	165600	0	0.00000
24	0	86400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	172800	172800	0	0.00000
25	0	90000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	180000	180000	0	0.00000
26	0	93600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	187200	187200	0	0.00000
27	0	97200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	194400	194400	0	0.00000
28	0	100800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	201600	201600	0	0.00000
29	0	104400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	208800	208800	0	0.00000
30	0	108000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	216000	216000	0	0.00000
31	0	111600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	223200	223200	0	0.00000
32	0	115200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	230400	230400	0	0.00000
33	0	118800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	237600	237600	0	0.00000
34	0	122400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	244800	244800	0	0.00000
35	0	126000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	252000	252000	0	0.00000
36	0	129600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	259200	259200	0	0.00000
37	0	133200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	266400	266400	0	0.00000
38	0	136800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	273600	273600	0	0.00000
39	0	140400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	280800	280800	0	0.00000
40	0	144000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	288000	288000	0	0.00000
41	0	147600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	295200	295200	0	0.00000
42	0	151200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	302400	302400	0	0.00000
43	0	154800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	309600	309600	0	0.00000
44	0	158400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	316800	316800	0	0.00000
45	0	162000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	324000	324000	0	0.00000
46	0	165600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	331200	331200	0	0.00000
47	0	169200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	338400	338400	0	0.00000
48	0	172800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	345600	345600	0	0.00000
49	0	176400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	352800	352800	0	0.00000
50	0	180000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	360000	360000	0	0.00000
51	0	183600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	367200	367200	0	0.00000
52	0	187200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	374400	374400	0	0.00000
53	0	190800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	381600	381600	0	0.00000
54	0	194400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	388800	388800	0	0.00000
55	0	198000	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	396000	396000	0	0.00000
56	0	201600	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	403200	403200	0	0.00000
57	0	205200	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	410400	410400	0	0.00000
58	0	208800	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	417600	417600	0	0.00000
59	0	212400	0.000E+00	0.000E+00	0.000E+00	0.000E+00	0.00000	1	1	0	-0.0001	0	0	0	0.00	0.00	424800	424		




Start of Outflow (h:min)
14602
1523146
151774
1519207
164

