REVISED PARAMETER ESTIMATION METHODS FOR THE PITMAN MONTHLY RAINFALL-RUNOFF MODEL

A thesis submitted in fulfilment
of the requirements for the degree of
Master of Science
of
Rhodes University
Grahamstown
South Africa

By

Evison Kapangaziwiri

October 2007

Dedication

This thesis is dedicated to my friend and wife, Concillia, and our son Hanyadzashe who had patiently endured my long absence from home. Thank you guys for all your understanding and support. Love ya!

Acknowledgements

I am grateful to Prof Denis Hughes, my supervisor, for the support and guidance throughout the conception and execution of this research work. I am especially thankful for his patience in going through the numerous manuscripts that preceded this final report. The equations developed for the estimation procedures were packaged into a small and neat computer program which he coded for me. Thank you also to David Forsyth for assisting with the conversion of data to formats acceptable in the SPATSIM software. A number of small programs were written for my benefit and this is greatly appreciated.

A special word of thanks goes out to the providers of data used in the development and testing of the estimation methods in this thesis, Department of Water Affairs and Forestry (DWAF), the Zimbabwe National Water Authority (ZINWA), Dr Rikard Liden of SWECO Consultants (Sweden), Mr. Anthony Mporokoso of Department of Water Affairs (Zambia). I also would like to express my gratitude to numerous colleagues who help shape this thesis through constructive comments and encouragement.

I would like to thank Institute for Water Research (IWR) for their financial support in the form of bursary for 2006 and a partial bursary for 2007. The financial assistance of the department of Labour (DST) towards this research is also hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DST or IWR.

I would like to thank all members of staff of the Institute for Water Research for various fruitful and enlightening discussions that helped my appreciation of various water related issues.

ABSTRACT

In recent years, increased demands have been placed on hydrologists to find the most effective methods of making predictions of hydrologic variables in ungauged basins. A huge part of the southern African region is ungauged and, in gauged basins, the extent to which observed flows represent natural flows is unknown, given unquantified upstream activities. The need to exploit water resources for social and economic development, considered in the light of water scarcity forecasts for the region, makes the reliable quantification of water resources a priority.

Contemporary approaches to the problem of hydrological prediction in ungauged basins in the region have relied heavily on calibration against a limited gauged streamflow database and somewhat subjective parameter regionalizations using areas of assumed hydrological similarity. The reliance of these approaches on limited historical records, often of dubious quality, introduces uncertainty in water resources decisions. Thus, it is necessary to develop methods of estimating model parameters that are less reliant on calibration.

This thesis addresses the question of whether physical basin properties and the role they play in runoff generation processes can be used directly in the estimation of parameter values of the Pitman monthly rainfall-runoff model. A physically-based approach to estimating the soil moisture accounting and runoff parameters of a conceptual, monthly time-step rainfall-runoff model is proposed. The study investigates the physical meaning of the model parameters, establishes linkages between parameter values and basin physical properties and develops relationships and equations for estimating the parameters taking into account the spatial and temporal scales used in typical model applications. The estimation

methods are then tested in selected gauged basins in southern Africa and the results of model simulations evaluated against historical observed flows.

The results of 71 basins chosen from the southern African region suggest that it is possible to directly estimate hydrologically relevant parameters for the Pitman model from physical basin attributes. For South Africa, the statistical and visual fit of the simulations using the revised parameters were at least as good as the current regional sets, albeit the parameter sets being different. In the other countries where no regionalized parameter sets currently exist, simulations were equally good.

The availability, within the southern African region, of the appropriate physical basin data and the disparities in the spatial scales and the levels of detail of the data currently available were identified as potential sources of uncertainty. GIS and remote sensing technologies and a widespread use of this revised approach are expected to facilitate access to these data.

Table Of Contents

List Of Figures

List Of Tables

1 INTRODUCTION

1.1 Background

The complexity of current water resource management poses many challenges. Water managers need to solve a range of interrelated water dilemmas, such as balancing water quantity and quality, flooding, drought, maintaining biodiversity and ecological functions and services. The reliable quantification of hydrological variables such as rainfall and streamflow is a prerequisite for sustainable water resource management, planning and development within basins. Southern Africa's hydrological regime is characterized by high variability and low runoff coefficients with less than 15% conversion of mean annual precipitation (MAP) to mean annual runoff (MAR) known to be present across large parts of the region (Walmsley, 1991). With predictions of water scarcity conditions, due to rapid population growth, expanding urbanization, increased economic development and climate change, being predicted for the region (Rosegrant and Perez, 1997), water looks set to become a limiting resource in Southern Africa. The dynamics of demand and supply will have a large impact on the future socio-economic development of the region (Basson et al. 1997; Rosegrant and Perez, 1997). The other huge problem in southern Africa is the trans-boundary nature of a number of the river systems (e.g. the Zambezi, Limpopo, Orange, Okavango). This makes decision making for both the present and the future very difficult and uncertain and it is imperative to create a common platform for the quantification of this precious resource.

It is therefore prudent to be able to quantify the water resource adequately for meaningful management decisions, not only for the present but also for the future. However, data paucity as a result of shrinking measurement networks due to economic and manpower problems (Hughes, 1997; Oyebande, 2001) has had a limiting effect. Some of the major river systems in the region have been gauged for the determination of hydrological variables, but this has not been the case with most medium and small sized basins. Even so, there are several major basins in different parts of the region that are not adequately gauged and in some basins the existing gauging networks are being discontinued; this leads to uncertainty in the design of water resource systems. However, in spite of these problems water resource developments must continue to take place to satisfy the economic and social development needs of communities (Mazvimavi, 2003). To alleviate the problem of data paucity, hydrological and ecological simulation models have been used extensively in the region and water resource planning has thus often been highly dependent on their results. The Pitman model (Pitman, 1973) is an example.

The Pitman model has grown to be a widely used hydrological assessment tool in the Southern African region and it is the author's conviction that it could be used to a greater extent in the future. Its simplicity and user-friendly interface make it an attractive option and its data requirements are quite simple and easily met by most of the region's hydro-meteorological agencies. The major limitation of the Pitman model is the number of model parameters that need to be optimized which often makes it harder to apply consistently in data scarce regions like southern Africa. However, with the current impetus in hydrology being the improvement of methods that enable hydrological predictions to be made in basins with limited or no historical measurement records and the reduction of the uncertainties associated with these predictions (Sivapalan, et al., 2003), the problem may be resolved. This study is borne out of the initiative of the International Association of Hydrological Sciences (IAHS) for predictions in ungauged (PUB).

The most popular of the traditional methods for prediction in ungauged basins has been the use of parameter regionalization. This involves the calibration of the model against naturalized observed flows and then developing statistical relationships between the parameters and some basin physical attributes or using some parameter mapping based on basin similarities. Two problems have always dogged regionalization in southern Africa – the limitations of flow monitoring networks mean there are generally limited reliable observed data for the calibration of the model and that there are quality issues with the data that are available. Unquantified upstream water use and abstractions mean that there are uncertainties with regards to the extent to which the observed flow data represent the natural hydrology of the basins. Given that there are other data collected by various agencies across the region (e.g. soil hydraulic properties, geology) that can be used to aid the understanding of the rainfall-runoff transfer process, this study therefore addresses the question of whether physical basin properties and the role they play in runoff generation can be used directly in the estimation of parameter values. If the answer is yes, then it may be possible to develop procedures for parameter estimation in ungauged basins that are less reliant on limited calibration results that are themselves likely to generate values with a degree of uncertainty.

There are a total of 41 parameters (only 11 are free/calibration parameters, the rest are estimated from basin properties) in the version of the model being used and the focus of this study is on the 7 calibration parameters that control the soil moisture accounting, runoff and recharge and the soil surface infiltration routines The prospect of 'free simulations' (using the model without calibration) would then be possible and could be used to generate flows in data scarce areas and ungauged basins. While there might be issues with the use of free simulations, they are definitely better than not having any information at all (Bergstrom, 1991). More robust parameter estimation procedures based on the physical basin characteristics may reduce the uncertainties associated with these.

1.2 Aims and Objectives

While the ultimate goal of a study of this nature would be to develop regional parameter sets for southern African basins similar to those established during the South African water resources assessment project of the 1990s (Midgley et al., 1994), the main aim of this study is to produce revised and improved calibration and application (in ungauged basins) procedures for the Pitman model in southern African basins under different climate, topography, geology, soils and vegetation conditions. This involves the estimation of parameters using conceptually physically sound principles which can be related to measurable basin characteristics and would be easier to evaluate in ungauged basins. A key goal in the development of the estimation procedures is to minimize the need for a basin-specific model calibration, and to achieve this, the model parameterization is to be structured around the use of basin physical and hydro-meteorological data.

To achieve this overall the following specific aims are envisaged for the study:

i. To develop a conceptual framework for the physical interpretation of the Pitman model parameters.
ii. To develop equations for the direct estimation of model parameters from physical basin property data.
iii. To generate sets of parameters for the Pitman model for selected basins in southern Africa.
iv. To assess the simulation results based on the use of revised estimation procedures in selected basins within southern Africa.

1.2.1 Developing a conceptual framework for the physical interpretation of the Pitman model parameters

Before physically-based estimation procedures can be developed for the Pitman model, it is essential to revisit the conceptual structure of the model and the way in which this relates to real hydrological processes. In doing this it is also necessary to consider the spatial and temporal scales at which the model is typically applied. To achieve this requires that the effect of each parameter be isolated and studied in depth to identify their physical meaning. This is what is meant by a conceptual framework for the interpretation of the parameters.

1.2.2 Developing equations for the direct estimation of model parameters from physical basin property data.

The conceptual framework will identify the specific hydrological response effects of each parameter. Using well understood principles of conceptual physical hydrology it should be possible to identify the physical basin properties that are relevant to individual parameters and develop equations that can be used to estimate the parameters. Once again, scale effects will need to be considered as will the typical availability of basin property data.

1.2.3 Generating sets of parameters for selected basins

Generating parameter sets for selected basins in the region requires the collection of appropriate basin property data. It was recognized at the start of the study that the sources, spatial resolution and accuracy of such data would vary considerably within the region, and clearly affect the results. However, this is part of the reality of applying estimation procedures in data scarce regions. Only the soil moisture accounting, recharge, runoff and soil surface infiltration parameters are being investigated and the other parameters would have to be calibrated where no regionalized parameter sets currently exist. For South Africa where regionalized parameter sets exist, those parameters estimated by the revised procedure will be used with existing parameters (not part of the new procedures) and with the same input data (rainfall and evaporation) used within the WR90 database (Midgley, et al., 1994).

1.2.4 Testing the parameters from revised estimation procedures in selected basins

The revised parameter set will be used in the model and the simulation results compared with observed flow data as well as previously established results using WR90 parameter sets in the case of South African basins. The results comparisons will be based upon a standard set of statistical criteria. One of the issues to consider is that few southern African observed flow data are completely natural, while it is often difficult to properly quantify the upstream development impacts. This issue will necessarily have to be considered in the selection of the test basins and in the interpretation of the results.

1.3 Research Questions

This study directly explores some of the principle issues associated with PUB, uncertainty analysis (though this part is not specifically covered in this study) and the practical application of hydrological models, in particular the Pitman model (Pitman, 1973), in a data scarce region such as southern Africa. The study will attempt to provide answers to the following questions:

i. How can we estimate hydrologically relevant model parameters?
ii. Can model parameters be defined in a physical manner that is consistent with physical hydrology principles?
iii. What are the optimal process conceptualizations for parameter estimations?
iv. What are the physical basin characteristics that affect model parameters?
v. What relationships exist between the parameters and the physical basin characteristics?
vi. What are the most robust ways of estimating parameters? Given the availability of data in southern Africa.
vii. What alternative sources of data can be used to aid the parameter estimation procedures?
viii. How can this knowledge be used to develop new guidelines for the calibration and application of the model?

It is accepted that models are generally quite simplified representations of reality whose parameter quantification is one of the major sources of uncertainty. This is especially so for conceptual models like the Pitman model. Therefore, in attempting to answer these questions a number of science issues would naturally arise and these would also need attention during the whole process. The issues that arise in this study are related to the following questions:

i. What are the limits of available data sets and what new data are required?
ii. Are there alternative conceptualizations (at the appropriate model spatial and temporal scales) of the natural hydrologic processes that will facilitate better parameter estimation procedures?
iii. What are the criteria for acceptability and are these sufficient?
iv. What is the uncertainty of using these estimation methods? How does this uncertainty propagate to the ungauged basin? What would the risk be in practice?

1.4 Expected research outputs and research justification

Given the regional situation, this study has the potential to provide a practical solution for water resource managers who are often called upon to make hydrological predictions in data scarce areas for long term, highly capitalized water resource development projects. It is realistic to believe that the Pitman model will remain to be a standard hydrological tool in the southern African region for a long time to come. The scope of use of the model will also continue to rise as more uses of the model are discovered. Chief among these may be the need to model water resource impacts of climate change. Published results indicate that climate model results at the monthly time resolution have been more reliable than at shorter time scales. Thus the Pitman model could possibly be used in the forecasting of water resource scenarios in analyzing and planning for the impacts of climate change (Hughes, 2004b). Simple, more objective and robust parameter estimation procedures can only be beneficial to the southern African community of water practitioners. The non-reliance of the proposed estimation methods on limited calibration results means that improved model regionalization could be achieved. In addition, a reduction in the subjectivity associated with traditional regionalization of model parameters could create greater common understanding across the region and foster improved relationships in trans-boundary river systems.

The study is expected to produce revised, physically-based estimation procedures for the soil moisture accounting, runoff, recharge and infiltration parameters of the Pitman model for some selected basins in southern Africa.

This desktop study has been designed to cover a number of selected basins in southern Africa. 71 basins were selected for the study and were chosen to span the range of basin physical and hydro-meteorological conditions obtaining in the region. In order to reach reliable conclusions it was necessary that the data be of reasonable length of at least 25 years, covering the hydrological and climatological regimes of any chosen basin. The data relevant for this study are monthly streamflow, precipitation and evapotransipration records. The data have been accessed from published information, on-line databases and through direct contact with the relevant data collection agencies. For the streamflow data it was considered appropriate to avoid using basins where the observed data are expected to be substantially impacted by upstream developments. Notwithstanding the time factor, naturalizing the flows demands the availability of records related to storage, abstractions or return flows which are frequently difficult to obtain in the region. Therefore, only basins with as near natural flow as possible, or with minimal human impact, would be chosen.

The remainder of this document consists of chapter 2 which contains a discussion of hydrological modelling with an emphasis on southern Africa. Chapter 3 gives a brief introduction to the Pitman model (Pitman, 1973) and its application as part of SPATSIM (Spatial and Time Series Information Modelling) framework software. This is followed by a detailed description of the model and its parameters which establishes the conceptual framework on which this study is based in chapter 4. A description of the developed parameter estimation procedures follows in chapter 5. Results are presented in chapter 6 with the discussion, conclusions and recommendations finalizing the report in chapter 7.

2 RAINFALL-RUNOFF MODELLING

2.1 Introduction

A model is a mathematical or physical analogue of a natural system (Linsley, 1981). It represents an abstraction of complex reality into a form that is more easily understood. Models are therefore simplified representations of the real system which is too complex to formulate in detail and their goal within the scientific community is to help understand the operation of and make predictions the real system (Corwin, 1996). The fundamental hydrological problem is the derivation of a relationship between basin rainfall and the resultant runoff. Hydrological modelling has its roots in the work of Pierre Perrault in 1674 whose endeavors to describe the relationship between basin runoff and rainfall resulted in a simple equation: Q = P/6, where Q and P were the annual runoff and precipitation respectively (Linsley, 1981). Other early hydrologists such as Edme Mariotte (1620 – 1684) and Edmond Halley (1608 – 1680) had almost the same relationship. Many developments have followed this pioneer work to the various models of varied description and complexity that are in operation today the world over. Hydrological modelling experienced a boom in the 1970s largely as a result of advances in computing technology.

A casual search for literature (e.g. ScienceDirect on Elsevier gives about 1200 articles) on hydrological modelling reveals many hundreds of papers covering a wide variety of approaches. There are those that focus on the hydrological understanding of the modelling process where physical hydrology principles drive the modelling process. Physical concepts are studied and understood before a decision on their adequate representation in a model is taken. Some deal exclusively with the mathematics of modelling in which emphasis is on such issues as the solutions to differential equations, optimization methods, objective functions, etc. The hydrological content is often very small, the focus being on mathematics. The last approach has been to deal essentially with the 'modelling' issues themselves where attention has been on the improvement of model efficiency, issues of uncertainty and equifinality (Beven and Binley, 1992) and the type of equations that models can use. With such a variety of approaches, it is a daunting task in the early 21st century to embark on a comprehensive review of the literature on hydrological modelling over the last four decades.

2.2 Model types and structure

This section attempts a description of the various model types and structures that have been used since the advent of modern day hydrological modelling. This discussion is not meant to be exhaustive but merely provide background to the current study. Models fall into many different classes. In an early treatise on hydrologic modelling, Clarke (1973) identifies two broad model classes. Models are classified on the basis of their description of the natural phenomenon into either regression or process-based models. Regression models recognize that hydrological events depend on chance and make use of historical hydro-meteorological time series data (e.g. rainfall and streamflow) and statistical principles to predict output in line with statistical patterns. On the other hand, process-based models use mathematical equations to describe hydrological phenomena in a particular basin based on the hydrological processes perceived to be in operation. The following is a small subset of process-based models; HBV (Bergstrom and Forsman, 1973), NAM (Nielsen and Hansen, 1973), Pitman (1973), TOPMODEL (Beven and Kirkby, 1979), ACRU (Schulze, 1986), IHACRES (Jakeman et al., 1990), VTI (Hughes and Sami, 1994), MIKE SHE (Refsgaard and Storm, 1995), Tank (Sugawara, 1995) and ARNO (Todini, 1996).

The process-based hydrological models can further be subdivided into either stochastic or deterministic in nature with the former assuming a randomness or uncertainty in the simulated output as a result of uncertainties in input variables (Beven, 2001). The latter relate to a simulation that allows only one outcome from given inputs on the assumption that processes can be defined in physical terms without a random component (Beven, 2001; Linsley, 1981) and are therefore chance independent (Clarke, 1973). Thus, the processes of transfer of rainfall to runoff are assumed to be governed by definite physical laws and a basin is a not a random assemblage of independent parts but an integrated physical system whose temporal and spatial variation can be adequately described (Pitman, 1973).

Deterministic models can further be classified as either empirical/metric or conceptual or physical. Empirical models are observation–oriented and characterize system response by extracting information from existing data (Kokkonen and Jakeman, 2001). This type of model is therefore essentially used to predict, but not explain, system function. Their development takes little or no cognizance of the features or processes of the hydrological system. There is no perceived attempt to understand the rainfall-runoff hydrological processes operating within a basin. The models require records of both river flow and rainfall for calibration, use curve fitting procedures and generally cannot be applied to the ungauged situation without modification. The roots of such models can be traced back to the unit hydrograph theory by Sherman in the 1930s which is based on the assumption of a linear relationship between rainfall excess and runoff. The Rational Method is another example of this type of model (Kokkonen, 2003). To reproduce the basin-wide streamflow response to climate inputs with an empirical model, it would suffice to have a lumped loss function, to account for processes such as evaporation, interception, surface and sub-surface moisture storage and groundwater recharge, and a routing function, to represent the different components of a basin's response (Wheater, 2005).

A physically-based process-oriented model is a simplified version of the real hydrologic system and tries to reproduce as much of the hydrological behaviour of a basin in the rainfall to runoff transfer process as possible. It is based on the generic understanding of the physics of the basin hydrological processes. Physical models recognize the entire basin as a spatially variable system and attempts to model a range of processes operating at small scales like a hill-slope to those in the entire basin. The basin is generally divided into a network of interlinked component segments for which all significant hydrological parameters are assumed to be measurable in the field. Complex mathematical relationships such as partial differential equations are normally solved numerically to describe the hydrological processes. The first of such models was developed in the 1970s in which finite difference methods were used to solve the Richards' equation for two dimensional unsaturated flow to represent slope hydrological processes. The Systeme Hydrologique Europeen (SHE) (Abbott et al, 1986; Bathurst, 1986) model is one of the well known physical models developed along similar lines. While these models may provide mathematically sound representation of hydrological physics, usually at smaller scales, they require comprehensive data and extensive computations. The models are, however, characterized by parameters that are, in principle, measurable and have a direct physical interpretation. Theoretically, therefore, if the parameters could be determined a priori, then such models could be applied to ungauged basins and the effects of basin climate variability or land use change explicitly represented. They should be appropriate for integrated basin modelling where considerations such as land use changes and/or climate variability, movements of pollutants and sediments and groundwater recharge are important outputs, e.g. the ACRU model (Schulze,

1995) and the SWAT model (Arnold and Allen, 1996). However, this has not been achieved in practice owing to the massive extent of the data demanded by such models and the simple fact that such data are not easily available or measured. The transfer of physics-based equations developed in the laboratory at very small scales to the typical modelling scale of a basin is not an easy task (Beven, 1989). The issue here is whether it would be possible to apply theoretical equations in the field, even at relatively small scales where there is extensive natural heterogeneity. This variation in physical basin properties is an important limitation in the wider application of physical models. Physically-based models are therefore more suitable for small scale research studies where the effects of basin heterogeneity and variability of parameters is small (Bergstrom, 1991).

Conceptual models represent a compromise between the two extreme modelling approaches outlined above. Conceptual models have proved to be the most common and parsimonious model type. Conceptual models describe all the component processes of the hydrological cycle considered important in a system. The natural hydrological system is represented as a system of interconnected storages, which would be recharged and depleted by appropriate component processes of the hydrological system. In this system moisture accounting of the input of rainfall is partitioned and routed to eventually produce runoff. The level of conceptualization of such a model reflects the extent to which the model structure and its parameters are representative of basin-scale hydrological processes. This approach is thus essentially semi-physical, where an understanding of the hydrological processes and process representation are integral to the modelling philosophy, but without comprehensive detail as in the physical models. The sizes of the moisture storages, moisture routing between these storages and the output of runoff are all described via mathematical equations (Nash and Sutcliffe, 1970). Besides the coefficients of the mathematical relationships, called parameters, which vary spatially and at times temporally, the core structure of the mathematical relationships is assumed to be constant for all basins. Beven (2001) and Corwin (1996) outline the components of the classical conceptual modelling process as model perception, conceptualization, verification, sensitivity analysis, calibration and validation. The conceptual model form became popular in the late 1960s and early 1970s thanks to computing power which then made possible an integrated approach to the land component of the hydrological cycle, albeit using simplified relationships, to generate continuous flow (Wheater, 2005). The application of this type of model to a basin usually requires the quantification of the parameters that describe the model in that particular basin and, for their extension to the ungauged basin, the development of relationships between the parameter sets and the basin physiographic characteristics. An attendant risk of conceptual modelling is that as the number of component processes increases so would the number of parameters and the uncertainties associated with their quantification. The Pitman (Pitman, 1973), the HBV (Bergstrom and Forsman, 1973), and the Sacramento (Burnash, et al., 1973) models are examples of the conceptual model type. Since their demand for input data is usually, though not always, minimal, conceptual models can be used in areas of data deficiency, making them practical and useful tools in operational hydrology. They are used mostly in hydrological forecasting (e.g. the operational flood forecasting system in the United states is based on the Sacramento model and in Sweden and Finland the HBV model is used), reservoir operation (e.g. HBV in Sweden) and in the extension of and filling in of gaps in observed records. They are also used extensively for a wide variety of water resource assessment studies (for example, Midgley et al., 1994).

Another distinct classification recognizes deterministic models as being either lumped or distributed. This classification is based on the spatial description of basin processes. A lumped model is one in which the parameters, inputs and outputs are spatially averaged for the whole basin. In a distributed model the basin is treated as a spatially variable system with all variables and parameters being allowed to vary spatially in response to differences in basin characteristics as well as rainfall and other climatic variables. There are two main groups of distribution system – one based on rectangular grids and the other on the use of natural drainage units. These drainage units vary in size from small hill-slopes or first order basins to larger basins such that the model network consists of lumped models connected by some routing system. Parameter quantification in lumped models usually requires that a historical observed flow record be available against which the model is calibrated. The parameters are averages over a basin and do not usually have any 'true' physical meaning. This unfortunately limits the applicability of such models beyond the areas for which they have been calibrated and their use in ungauged basins is problematic. However, a concise and unambiguous physical interpretation of the parameters should extend their applicability. In a fully physically-based distributed model attempts are made to infer parameter values from measurable physical and hydro-meteorological basin properties, thus rendering calibration unnecessary when sufficient data are available. In reality, however, some calibration is often necessary as it is impossible to characterize all the spatial and temporal variability of the parameters at the basin scale. This is the typical model scale for most practical purposes in water resource assessments and planning and is generally larger than the design scale of distributed models. Thus, some lumping is inevitable and the parameters become spatial averages rather than direct point values from field measurements. The 'probability distributed' approach of Moore (1985), who suggested that 'sub-grid' or 'sub-basin' effects could be accounted for through probability distribution functions representing the (largely unknown) variations in process functioning (and therefore parameter values) within a spatial unit, could be used. The parameters of a distributed model can, at least in theory, be validated by field measurements. This should make them more reliable to use in ungauged basins. In general, however, the large data requirements and structural complexity of distributed models make them less favorable for routine use than their normally parsimonious lumped counterparts.

Hughes (2004b) considers a classification based on model complexity. Model complexity is envisaged as "the extent to which the model attempts to represent the many and diverse processes that affect the response of runoff to rainfall". More complex models would attempt to explicitly represent all hydrologic processes (of interception, infiltration, soil water drainage, evapotranspiration, groundwater movement, etc) in a basin. This inevitably means more parameter space and more time, effort and information required in order for them to be used with any degree of success since all parameters would have to be quantified. The fully distributed process-based model types are a good example of the more complex type. Three methods of classifying models, spatial complexity, temporal complexity and model purpose, are identified. A "spatially complex" model is one in which the total basin is disaggregated into a number of sub-basins based on natural drainage units (e.g. slopes, channels, basins) or on geometric shapes (square grids, polygons, etc). The rationale of adopting this approach is that the parameters and the variations in input climatic data may be realistically represented, with limited heterogeneity, within smaller units of the total basin.

"Temporal complexity" groups models based on the time-step used. Time steps can vary from coarse intervals of a month or more, to fractions of an hour and to variable time scales (Hughes and Sami, 1994; Hughes, 1993a). With smaller time intervals, it is possible to simulate more realistically and in greater detail such rapidly changing hydrologic responses as floods. The two classes discussed above are not necessarily independent of each other and the separation in the discussion has only been for convenience. For each model structure a choice is made about the methods appropriate for the representation of the hydrologic processes and this decision necessarily influences the time interval of modelling and the spatial distribution system to be used. Lastly, "modelling purpose" groups models according to the type of outputs generated. While some models simulate single events only, there are a number that are designed to be multi-purpose and generate a wide array of continuous information such as moisture status, groundwater levels, channel transmission losses and recharge (see Hughes and Parsons, 2005).

Even though more complex models use more input data and have a more detailed process description there is no obvious relationship between model complexity and the quality of simulation results (WMO, 1975). In fact the more complex a model is, the more likely the problem of over-parameterization and the attendant parameter interdependencies. Such structurally complex and distributed models have been useful for research and process investigation and understanding. Simple, general water balance models can work fairly well for most practical hydrological purposes even though they may not represent the physical operation of the basin in a lot of detail. Simple models therefore have a role to play in the field of water resource estimation and increasing their complexity could be counterproductive.

Notwithstanding all the various model classes, the distinction between model types is not always obvious and models firmly rooted in one approach may exhibit characteristics of a different type. In reality, therefore, almost all models are usually crossbreeds from a number of model formulations and philosophies. Many models are therefore classified as, but not limited to, lumped conceptual models, distributed physically-based or, process-oriented or semi-distributed conceptual models. The Sacramento (Burnash et al., 1973) and the Stanford Watershed model (Crawford and Linsley, 1966) are common examples of lumped conceptual models. The Pitman model (Pitman, 1993) and the HBV (Bergstrom and Forsman, 1973) are examples of semi-distributed conceptual models while SHE (Refsgaard and Storm, 1995) is a common example of a fully distributed physically-based model.

2.3 Parameter interdependence and sensitivity

Though models do vary in complexity and structure, nearly all have parameters for which values must be somehow quantified. A parameter is a quantity that characterizes an aspect of a hydrological system in a particular basin and should remain constant in time. Parameters are distinct from variables in a hydrological system which are measurable characteristics of the system that assume different numerical values at different times (Clarke, 1973). Thus input rainfall, simulated soil moisture and both observed as well as simulated runoff are model variables. The number of parameters in a model has often been used to determine its level of parsimony as there is usually a positive correspondence between model complexity and the number of parameters. Parameters are an inherent component of all models and are basin or sub-basin specific with some having been observed to vary seasonally and still others being dependent on the spatial or temporal scales used.

Within any model, parameters exhibit elements of interdependence with each other, the extent depending on the structure and complexity of the model. Parameter interdependencies make the process of model calibration very difficult. A parameter response surface represents the value of one or more objective functions associated with results of varying two or more parameter values. Calibration is aimed at identifying the optimal location on this surface (either maximum or minimum, depending on the objective function). An objective function is a statistical function associated with an optimization problem and determines the success of a solution. It measures the match between simulated and observed time series. (see section 2.4.3). Calibration of the parameters, especially by an automatic algorithm, is thus often referred to as hill climbing in reference to the progressive attempts to get to the optimal solution. Generally optimization algorithms are categorized as either local or global where the later are designed to locate the global optimum and not be trapped at a local optimum (Madsen, 2000). Popular stochastic global search criteria include the shuffled complex evolution (SCE) algorithm (Duan et al., 1992) and genetic algorithms (GA) (Ndiritu and Daniell, 2001). The SCE has been used extensively in the calibration of conceptual rainfall-runoff models e.g. the Sacramento, NAM, Xinanjiang and the Pitman models (Gan et al., 1997) and physically-based models (see Duan et al., 1992). It has proved to be a reliable and efficient automatic optimization tool.

The interdependence characteristic of parameters has led to problems of parameter unidentifiability, over-parameterization and equifinality (Beven, 1993; 2001). A parameter is said to be unidentifiable if it cannot be estimated from a given data set. It is theoretically unidentifiable if it can never be estimated, no matter how extensive the data set is. If the best value of a parameter depends on the values of other parameters, then it is not identifiable. On the other hand a model is said to be over-parameterized if there are too many degrees of freedom in relation to the amount of information that is contained in the observed record. Wheater et al. (2005) suggest that on the basis of information normally available in observed time series flow data, a maximum of five or six parameters would be adequate to sufficiently describe a system. This, however, seems to be an arbitrary mathematical interpretation. In practice, if models are to be used in a more physically-based manner in ungauged basin it is often necessary to have increased parameters in order to adequately describe the basin response to meteorological input. While fewer parameters may be attractive, their physical relevance is highly compromised. Calibration is then reduced to optimizing an objective function ignoring any connection between parameter values and basin properties. This problem is exacerbated in regional model calibrations where consistent and unambiguous relationships, which can be transferred to ungauged basins, are one of the objectives. Equifinality defines the existence of a number of different equally good parameter sets within a given model structure that may be acceptable in the reproduction of the observed behavior of that system (Beven and Freer, 2001). This clearly is in contradiction with the traditional concept of model calibration which is, implicitly, built on the hypothesis of the existence of a unique parameter set. While these problems are not unique to any one modelling philosophy they are more pronounced in the conceptual type of model. Normally, at the basin scale, the conceptual model approach is rather crude and almost statistical in nature, based on the assumption that the spatial variability of physical processes is not adequately known. Therefore their parameters are reduced to being just averages over a large area which often is an integral of several processes and their variability (Bergstrom, 1991). This consequently makes the physical interpretation of the parameters quite vague and difficult. Consequently the parameter response surface of conceptual models is dotted with numerous local optima making the use of local optimization search methods ineffective since the estimated optimum solution will depend on the starting point of the calibration process. The usually unknown interactions of the parameters make the parameter estimation procedure and the regionalization of parameters very difficult.

2.4 Model parameter calibration and validation

2.4.1 Model Calibration

A model structure, through its parameters, needs to be established in order to adequately simulate the hydrologic response of a specific basin to meteorological inputs. In this process the parameters are continuously adjusted until the simulated time series is a reasonable match to the observed time series data. The process of adjusting parameters to get an optimal parameter set is known as calibration. 'Free' parameters are those that cannot be quantified from experience, from basin property data or from other sources of information. Calibration is a necessary step in hydrological modelling, regardless of the number of parameters and the complexity of the model structure. This is because most model parameters cannot be measured, which frequently is a consequence of ambiguous physical meaning (Ao, et al, 2006). Madsen (2000) outlines the following synthesis of the objectives, in operational terms, of calibration:

i. A good water balance shown by a good agreement between the average runoff volumes.
ii. A reasonable agreement in the shapes of the simulated and observed hydrographs.
iii. A reasonable agreement in the timing, rate and volume of the peak discharges.
iv. A reasonable reproduction of the observed low flows.

What can be considered 'reasonable' would, however, need to be quantified for an individual basin and may depend to a certain extent on the quality of the input data and the objectives of the modelling exercise. For a successful calibration the observed record must contain sufficient signals to guide the process. While there are no hard and fast rules about the length of calibration period, it is necessary for the data to be long enough to cover the spectrum of significant events experienced in the basin (wet and dry periods, for example). The period should cover as many signals as possible and also be long enough to establish a stable parameter set. The length of record required may vary between climatic zones (Görgens, 1984). Model calibration can be achieved manually or automatically.

2.4.1.1 Manual calibration

The manual calibration technique is the traditional and used to be the more widely preferred of the two techniques. It involves the manual (or 'expert') adjustment of parameter values to improve the model response, based on visual inspection of the observed and simulated hydrographs and assessment using statistical measures of correspondence or objective functions. The aim is to reproduce the hydrograph peaks (amount and timing), runoff volumes, recession slopes and baseflow. A successful manual calibration calls upon the experience of the modeller and an intimate knowledge of the basic processes and interactions in the model. Thus the process can be slow, laborious and frustrating, particularly when there are many parameters to optimize and many unknown interactions between these parameters. Manual calibration can be highly subjective. However, it is the only feasible approach in areas of data scarcity (like southern Africa) where it is possible that the use of automatic calibration may lead to an optimization against inadequate signals or errors in the data. While optimization could be achieved it may be for the wrong reasons and the parameter set may not be hydrologically sensible due to calibration against errors in the data set. Manual calibration is also useful in regional calibration where it is necessary to ensure that similar basins have similar parameters, so that guidelines are developed for the use of the model in ungauged basins.

2.4.1.2 Automatic calibration

To circumvent the apparent disadvantages of the manual process, automatic techniques were developed. Automated calibration is based on optimization theory and requires the definition of a statistical measure of the differences between the simulated and observed hydrographs (i.e. objective function) and uses a mathematical algorithm to search the parameter response surface for the optimum parameter set. Also required to enable an automatic search is an observed time series against which model performance is assessed and a termination criterion to stop the iterations (Madsen, 2000). Automatic calibration has developed rapidly since the early pioneer work by Dawdy and O'Donnell (1965), Nash and Sutcliffe (1970), Ibbitt (1970) and others. The motivation for automatic techniques has been: a) the need to speed up and simplify the calibration process, b) the need to assign some objectivity and confidence to the naturally subjective manual calibrations and hence, model predictions and c) the lack of numerous expert calibrators (Hogue, et al., 2000). The Generalized

Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) is an example of generic algorithms for automatic calibration of hydrologic models that has enjoyed widespread use.

The advantages forwarded in support of this method are that the computer, rather than the modeller, does the hard work of exploring the parameter space and that the procedure is relatively objective and would (at least in theory) provide, after thorough searching, a single optimum parameter set. Experience has, however, tended to disprove the later as no unique optimum parameter sets have been obtained for models (Beven, 1989; 1993; Beven and Freer, 2001). This has been attributed to the imperfect input data such as rainfall and evapotranspiration, over-parameterization or the fact that parameters, even for the physically-based process-oriented models, are only averages over heterogeneous landscapes (Wheater et al., 2005; Bergstrom, 1991). Wheater et al. (2005) further outline a number of issues that arise in automatic calibration as:

i. The existence of many local optima in the parameter space.
ii. Many known and unknown interactions among model parameters giving rise to problematic valleys, ridges and/or saddle points on the parameter response surface.
iii. Some parameters are insensitive beyond certain threshold values
iv. Scale issues e.g. where different parameters are determined at different scales making it difficult to define an optimization step size in each parameter direction when used simultaneously in a search of the response surface.

In many cases though, it is practical to use both calibration techniques to complement each other. Frequently in some conceptual models, it is wise to guide the calibration process by first roughly calibrating the model manually in order to get an acceptable, hydrologically sensible parameter set. Parameter boundaries would then be designed to constrain the subsequent automatic calibration. This ensures that the model simulations are hydrologically plausible and the simulation is for the right reasons.

2.4.2 Parameter validation

For the purposes of testing the adequacy of a calibrated parameter set in a gauged basin, it is necessary to perform a validation of the set on an independent period of the observed record. Parameter validation involves the comparison of the model output to the observed time series data with no adjustment or modification of the parameter set. Typically, validation performance statistics are worse than for the calibration period and if they are significantly lower, then questions about over-parameterization could be raised as the model might have many more degrees of freedom beyond the information contained in the observed calibration record. However, the calibration data set may also not have been representative. Klemes (1986) outlines some of the different ways of achieving a validation as;

i. Split-sample test - this is the use of two mutually exclusive subsets of the observed record.
ii. Differential split-sample test - used when the model is to simulate hydrologic response to climate and/or land-use conditions that may be significantly different from those of the available flow record, e.g. if a model was intended to simulate streamflow response to dry climate conditions it would be calibrated on a wet period and validated on the dry period.
iii. Proxy-basin test - relates to the use of one or several basins for calibration and validation in another, but homogeneous, basin. This is done, for instance, when a basin, with insufficient streamflow data, is to have its record extended. Adjustment of parameters on the basis of basin properties is allowable but not calibration.
iv. Differential proxy-basin test - is a combination of split-sample, differential split-sample and proxy basin validation. This is used for a general model aimed at accommodating all possible spatial variations.

Klemes (1986) further maintains that a model ought to be validated for the specific need and the types of application for which it is intended. Beyond these the model performance cannot be guaranteed and uncertainties are high. Besides Klemes' methods, in multi-purpose models, designed to produce a wider range of outputs, calibration and validation may be performed on two different output variables. Such multi-criterion validation assesses the goodness of simulations of different variables when the model is calibrated with respect to another variable

(Hughes, 1993b). This should reduce the problem of equifinality and help in the assessment of parameter and model uncertainty (Wheater, 2005) and test model stability. A validation test can also be performed based on regionalization where the calibrated optimum parameter values are related to basin characteristics. This kind of validation assesses the physical soundness of both the model and the parameter estimation procedures (Seibert, 1999).

2.4.3 Assessment of model performance

Regardless of which method of calibration is used, there is always a need to assess the performance of the model in any particular basin that is being modelled. This is achieved by measuring the extent to which the simulated runoff matches the observed runoff time series. Besides a visual inspection of the two time series hydrographs, usually associated with manual calibration, more objective statistical measures are also used. A statistical method, referred to as an 'objective function', is normally used to objectively assess the correspondence between the two time series. The aim of the calibration process, therefore, is to optimize (either minimize or maximize depending on the type of statistical measure being used) this objective function. There is a wide variety of objective functions cited in the literature and a modelling application usually determines the ones to use. All methods aggregate the time series of the residual errors over the whole modelling period. Given that there is so much information that can be obtained from an observed flow time series, it is not possible for all the different flow components (e.g. peaks, low flows, and recessions) of the data to be sufficiently evaluated by a single performance criterion. For a complete assessment, a number of objective criteria should be used. While a more comprehensive list of objective functions can be found in Görgens (1983), a small sample of common objective functions is listed here:

(i) Coefficient of Efficiency (CE): This is the Nash and Sutcliffe (1970) model efficiency criterion. The model efficiency has become one of the most widely used measures of goodness-of-fit in hydrological modelling. CE is a dimensionless relative index of correspondence between the simulated and observed time series. Comparisons of performance can be done over different periods or basins owing to the dimensionless characteristic of CE. It is given mathematically as:

$$CE = 1 - \left[\Sigma (Q_{obs} - Q_{sim})^2 / \Sigma (Q_{obs} - \bar{Q}_{obs})^2 \right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 2.1$$

where Q_{obs} is the observed time series, Q_{sim} the simulated time series and $\bar{Q}_{obs}$ is the mean of the observed series. CE can assume any values between -∞ and 1 with the latter indicating a perfect fit between the observed and the simulated flows. When CE takes the value of zero, the simulated flow is no better estimator than the mean of the observed flows and a negative value indicates that the simulated flow is a worse estimator than the mean observed flow. CE has been observed to give relatively high values even for some visually poor simulations. It is also difficult to get high CE values in basins or periods where the variation of streamflow is low. The value of CE is sensitive to systematic errors.

(ii) Coefficient of determination, R^2: relates to the proportion of variability within an observed time series data set that is explained by the simulated one. It is written mathematically as:

$$R^2 = \Sigma[(Q_{obs} - \bar{Q}_{obs})^*(Q_{sim} - \bar{Q}_{sim})]^2/[\Sigma (Q_{obs} - \bar{Q}_{obs})^2 * (Q_{sim} - \bar{Q}_{sim})^2] \ldots\ldots. \quad 2.2$$

where $\bar{Q}_{sim}$ is the mean of the simulated time series. R^2 varies between 0 and 1 inclusive and $R^2 = 1$ indicates that the simulated time series explains all variability in the observed time series, while $R^2 = 0$ indicates a poor correspondence between the two time series. While the CE is sensitive to systematic errors (general over- or under-estimation), R^2 is not similarly affected and a value close to 1 does not necessarily imply a good simulation. Where both the CE and R^2 are used as assessment criteria, large differences between them indicate systematic errors.

(iii) Percentage error of the total discharge volume (%V) or peak discharge (%P): these measure the percentage deviation in the total volume and peak discharge of the simulated from the observed. A perfect correspondence between the hydrographs of simulated and observed flows is shown by a value of zero with poor simulations being shown by an increasing divergence (in both directions) from zero. High values of %P and %V are an indication of systematic error. Low values of %P and %V can indicate low CE or R^2 values. The percentage error of total discharge volume is written as:

$$\%V = 100 * (VQ_{obs} - VQ_{sim})/ VQ_{obs} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \quad 2.3$$

where VQ_{obs} and VQ_{sim} relate to volume of observed and simulated time series respectively. A percentage error of the mean annual runoff (MAR) can also be used and is given by:

$$\%Mean = 100 * (MAR_{obs} - MAR_{sim}) / MAR_{obs} \quad \text{.......} \quad 2.4$$

For the %P, a threshold is defined above which all flows are regarded as peak flow events and it is expressed as;

$$\%P = (100/Y) * \sum (PQ_{obs} - PQ_{sim}) / PQ_{obs} \quad \text{.......} \quad 2.5$$

where Y is the number of peak flow events, PQ_{obs} and PQ_{sim} relate to the peak events of the observed and simulated time series.

(iv) Comparison of flow duration curves: A streamflow duration curve illustrates the relationship between the frequency and magnitude of streamflow and is a cumulative frequency curve that shows the percentage of time that specified discharges are equaled or exceeded. The flow duration curves of the simulated can be compared to that of the observed flow to judge the ability of the model to reproduce the flow pattern. Duration curves reflect the flow regime of the basin, with ranges from the low to the high flows being shown. This is a more reliable method for water resource assessments for the design of reservoirs or establishment of abstraction works. It is often quite possible to have a poor regional representation of rainfall in a basin which may fail to capture short-term variations. This would give unrepresentative simulated runoff time series, even without systematic error in the input data. In such cases it is, however, possible to get a representative duration curve as long as the model represents the rainfall-runoff response correctly.

All the objective functions can be calculated using untransformed (natural) streamflow data or using the natural logarithm transformed data. The logarithmic transformation of data removes the bias towards the high flow values and gives greater prominence to the moderate to low flows.

2.5 Modelling in ungauged basins

2.5.1 The Ungauged Problem

The discussions in the preceding sections have implicitly highlighted the importance of observed data in water resource assessments and studies. Frequently, however, historical streamflow records may not be available for various reasons such as a lack of gauging equipment or when an assessment is needed beyond the gauged circumstances, e.g. flood predictions, hydrological impacts of future land use or climate changes. Both cases, though in different ways, essentially represent an ungauged scenario.

By definition, an ungauged basin is one with inadequate hydrological observations, in terms of both data quantity and quality, to enable a computation of hydrological variables, at appropriate spatial and temporal scales, at a level of accuracy acceptable for practical water resource management (Sivapalan et al., 2003). These hydrological variables could include evaporation, infiltration, rainfall, runoff, or groundwater recharge and flow. However, many processes that are of interest in hydrology are usually difficult to observe routinely. Streamflow measurement is a variable that can be measured with considerable confidence at a gauging station in a basin and is the most important variable for most water resource planning purposes. Hence, from a practical point of view, the definition for an ungauged basin has, quite understandably, been reduced to refer to those basins with inadequate streamflow measuring facilities or those with scanty or no streamflow records. Strictly then, this makes almost every basin, to some extent, an ungauged basin. Important or strategic river basins in many parts of the world (including the southern African region) may have sufficient hydrometric stations for the determination of streamflow and other variables but many small to medium sized basins are ungauged. However, there are several major basins in different parts of southern Africa, for example, that have not yet been adequately gauged. In some basins the existing gauging networks are being discontinued, mainly due to economic and political constraints, both past and present, (see Hughes, 1997; Obebande, 2001). This has made large parts of southern Africa virtually ungauged. Unfortunately, water-related developments such as dam construction, irrigation development, etc. still have to take place in such data scarce situations and hydrologists are called upon to generate realistic water resource information. The problem of the ungauged basin is not a new one. Nash and Sutcliffe (1970) once remarked that "few hydrologists would confidently compute the discharge hydrograph from rainfall data and the physical description of the catchment" and that "this is a practical problem" (pp. 282) that hydrologists face in the field. The International Association of Hydrological Sciences (IAHS) has adopted 2003 – 2012 as a Predictions in Ungauged Basins (PUB) decade which is aimed at identifying a major breakthrough in the theoretical foundations of modelling and a critical examination of the existing approaches to hydrological predictions. The main objective of PUB is to move away from calibration based modelling towards "understanding-based methods" that would make predictions in ungauged basins simpler and less uncertain.

The primary cause of difficulties with hydrological predictions in ungauged basins is the high degree of heterogeneity of the land surface conditions such as soils, vegetation and land-use, as well as the space-time variability of the model inputs. In spite of many advances in developing methods for delineating homogeneous and heterogeneous regions on the basis of specific hydrologic variables (Nathan and McMahon, 1990), there remain numerous basic problems (inadequate measurement techniques, the availability of input data, scale issues, etc.) that still need to be resolved before these methods can be used universally. One of the drawbacks is that the existing methods for estimating the degree of heterogeneity require field data at regional scales, which are generally not available. Even in cases where sufficient data are available, representation of hydrologic processes resulting from the heterogeneities of landscape properties, land use and climatic change is difficult. There is also a need to understand more fully the sub-surface hydrological processes that are also important in the total rainfall-runoff transfer process.

2.5.2 Parameter Regionalization

For ungauged basins, the problem of model and parameter uncertainty is even more acute as no data are available to constrain predictive uncertainty. The most common and favoured way used to make hydrologic predictions in ungauged basins has been the extrapolation of information from gauged basins. This process is commonly known as regionalization (Nathan and McMahon, 1990). The basic tenet in regionalization is that, if there exists a relationship between model parameters and basin properties which holds for a gauged basin then flow simulations can be achieved in an ungauged basin which has similar physical attributes. The most common basin attributes that have been used include climate, topography, vegetation, soil properties (e.g. Chiew and Siriwardena, 2005), annual rainfall, areal potential evapotranspiration (e.g. Boughton and Chiew, 2006), basin area and geology. There are various means by which regionalization of models is achieved but they all tend to fall into one of the following groups:

(i) Statistical methods: Regression relationships are developed between optimized model parameters and some basin attributes for a number of gauged basins. Frequently, bivariate and multivariate linear and non-linear regressions are developed and then transferred to the ungauged basin (e.g. Boughton and Chiew, 2006). One of the weaknesses of such regression-based approaches is that the calibration parameters are subject to uncertainty and may have strong interactions among them, making them quite unstable. The calibrated parameters may reflect input data errors as well as true signals related to variations in basin properties. Parameter interdependencies and non-unique parameter sets suggest that the calibrated parameters may be partly a reflection of the calibration approach and contain some degree of subjectivity. Also some of the parameters may not be easy to estimate as a result of a lack of a concise physical interpretation. Hughes (1982) explored the transferability of conceptual model parameters for basins in South Africa and developed a calibration procedure which takes cognizance of the need for some physical interpretation of the parameters.

(ii) Parameter mapping: The simplest way to achieve regionalization would be to fix model parameters to average values for the region. This might be successful if the whole region exhibited the same hydrological response to rainfall input. Using a parameter set (either predetermined parameter sets or regional averages) for over 300 Austrian basins yielded very poor results (Merz and Bloschl, 2004). A more promising approach would be to assign a priori values to parameters based on some sort of similarity measure of basins using soils, rainfall, runoff ratios, etc. This sort of regionalization relies heavily on the premise of hydrologic similarity between some gauged and the ungauged basins and therefore the delimitation of hydrological response units (HRUs) based on chosen group-defining signatures (Nathan and McMahon, 1990). While the simplest way to define HRUs has been to use geographical proximity, it is not always a reliable of hydrologic homogeneity. This assumes that basins that are in close proximity to each other would have a similar runoff regime since climate and basin properties vary smoothly in space. Merz and Bloschl (2004) reported that the use of average values of parameters of immediate upstream and downstream neighbours gave the best results out of a suite of methods evaluated. Mazvimavi (2003) used cluster analysis to define hydrological response groups based on hydro-meteorological and other physical basin characteristics for Zimbabwean basins for the regionalization of the 'abcd' and Pitman models. Approaches such as the use of Kriging, neural networks and region of influence are also possibilities for regionalization studies (Fernandez et al., 2000). The parameterization of the so called quaternary basins in South Africa (Midgley et al., 1994) and the regionalization of the HBV in Sweden (Bergstrom, 2006) were both achieved by mapping parameters from gauged basin on the basis of similar hydrological response between the gauged and the ungauged basins.

The traditional approach to all regional calibration has been to use a two step method where firstly, optimal parameter sets are calibrated at individual gauged sites in a region of interest. Next, regional relationships between basin attributes and the parameter sets are developed and applied in the ungauged basins. The implication is that the parameter calibration process and the parameter regionalization process are treated independently. Fernandez et al. (2000) describe an alternative approach which they call 'regional calibration'. Instead of the twin optimization procedure, the regional approach proposes to simultaneously optimize the model parameter calibration and the regional relationships. The aim is thus to minimize model residuals and maximize the goodness-of-fit between model parameters and basin attributes at the same time. An application of the method to some 33 basins in the United States gave impressive results for the overall regional relationships but there was no significant improvement over the traditional method with regards the results at ungauged sites.

More recently, Yadav et at. (2007) have presented a new approach based on the regionalizing of relationships between catchment dynamic response behaviour and basin structure and climate. This approach deviates from the ones discussed in this section in that basin hydrologic response behaviour is estimated and regionalized within an uncertainty framework. The strength of this approach is in the model-independent quantification of the hydrologic behaviour through the use of streamflow indices and constraining expected catchment behaviour at ungauged stations.

(iii) *A priori* estimation methods: This is where the parameter values are fixed at values based on experience or the use of values adopted from the literature for basin characteristics. The Model Parameter Estimation Experiment (MOPEX) investigated the relationships between physical and hydro-meteorological basin attributes and the parameters of a number of selected hydrologic models. The intention was to develop enhanced a priori parameter estimation methods in a number of locations that were chosen to span several hydrologic, topographic and climatic regimes (Wagener et al., 2006; Ao et al., 2006). A huge database of basin characteristics and historical hydro-meteorological data was developed, and is being continuously expanded to incorporate as wide a spectrum of variations as possible. However huge this project may be, the methods are no different from the ones discussed in the preceding paragraphs.

It is important to note that testing regionalization approaches involves reserving some of the gauged basins to test the regional parameter estimations. This means that the data set used to establish the regionalization is reduced in size. This can be a problem in areas with a limited number of gauged basins such as southern Africa. Many regionalization studies have met with limited success (Franks, 2002). The problems that seem to haunt all the studies are equifinality and parameter interdependence. It has not been easy with most regionalization methods to be sufficiently confident that all the necessary and dominant controls of basin behaviour have been captured in the regionalization process. The regression equations derived from perceived relationships between basin characteristics and model parameters used are naturally empirical and therefore cannot be expected to be universally applicable, even in apparently similar basins. Significant bias exists in calibrated parameters due to observation error and model process uncertainty that permeate the derived regionalization techniques. This hinders the derivation of robust relationships on which ungauged basins can be confidently parameterized. This subsequently leads to high predictive uncertainty for ungauged basins (Franks, 2002). Ao et al. (2006) suggest that five aspects are required to achieve reliable regionalization based on direct parameter estimation using basin characteristics. These are model parameters that have exact physical meanings, a large amount of spatial physical property data, establishment of relationships between basin property data and parameter values, establishment of parameter-basin characteristic transfer functions and use of GIS techniques. Direct estimation of parameter values from physical basin attributes is more desirable in that it reduces the ambiguity in parameter estimation by calibration based only on the runoff signal at the basin outlet. For instance two different basins display a similar runoff pattern but have different basin attributes. Regional relationships and parameters developed using such basins are unlikely to be optimum, stable or reliable.

2.6 Parameter uncertainty in hydrological modelling

The discussions in some of the preceding sections have alluded to the issue of uncertainty in hydrologic modelling. Though not a part of this current study, a brief reference to this issue is imperative especially as it relates to parameter uncertainty and how it is dealt with in the discipline. Given that models are quite imperfect representations of the real hydrological processes, it is important to address the problem of uncertainty. Only then can decision makers evaluate how reliable simulation results are for the purposes of water resources management decision support (Kokkonen, 2003). Uncertainty in model parameterization and, subsequently, regionalization, stems from a number of sources including observation errors in input basin attribute data (e.g. runoff, precipitation, evapotranspiration, slope, infiltration, soil depth, etc), scale issues (i.e. the averaging of point measurements to represent larger areas) linked to both input data and parameters, insufficient attention to and poor understanding of some of the appropriate controls of basin response (e.g. subsurface moisture movement), model structure leading to process uncertainty, initial parameter ranges, choice of objective function, equifinality and the empiricism linked to most relationships between model parameters and basin physical properties (Ao et al., 2006). This often leads to uncertainty in model parameters, due to a lack of identifiability, which may limit significantly the use of models for such purposes as parameter regionalization or making predictions beyond the gauged circumstances, such as generating land-use or climate change scenarios.

Dealing with uncertainty in hydrological modelling is not easy and it can prove to be computationally demanding to assess its extent and effect on model results (Kokkonen, 2003). Uncertainty has important implications for decision making in water resources management and planning. One simple way of dealing with uncertainty would be to design less complicated, parsimonious model structures. However, caution needs to be exercised in choosing the number of processes to be represented as too simple a model structure may be impossible to use outside the range of conditions for which it was calibrated (Wheater, 2005). Another way to counter parameter uncertainty is to increase the amount of information available to identify the parameters, e.g. increasing the number of output variables. The success of this depends on the ability of the model structure to handle this extra load (Wheater, 2005; Beven, 2001). On the other hand the improved use of information already available to improve parameter identifiability is another alternative. For instance, different periods can be used to identify different parameters. This represents a multi-objective calibration approach for estimating model parameter values and evaluating model structural deficiencies (Wagener et al., 2001). Some algorithms have also been developed to deal with parameter identification and uncertainty estimation in hydrological models, e.g. Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992; Beven, 2001). This algorithm, apparently, suffers from an inefficiency to extract information from observed data and uses a definition of predictive uncertainty that is at variance with the widely accepted one (Mantovan and Todini, 2006). The problem of predictive uncertainty requires more research in order to get a solution.

2.7 Use of hydrological models in southern Africa

It has been just over three decades since a model designed for use in climatic conditions prevalent in most southern African countries was developed through the pioneer work of Pitman in 1973 at the University of Witwatersrand, South Africa (see Hughes, 2004b). Through different versions (see e.g. Pitman, 1973; Hughes, 1997; Hughes and Parsons, 2005) this model has been the most widely used in the region. It has been used for regional studies in the Okavango basin (covering Angola, Namibia and Botswana - see Hughes et al. 2006) and in selected basins of all countries in the region in the Flow Regimes from International Experimental Network Data (FRIEND) project (Hughes, 1997), water resource assessment studies in the Pungwe basin (which covers Mozambique and Zimbabwe) (SWECO, 2004) and the Kafue basin in Zambia (Mwelwa, 2004), for estimation of hydrologic variables and regionalization studies in Zimbabwe (Mazvimavi, 2003) and for simulation of arid climatic conditions in Namibia (Hughes and Meltzer, 1998) and Botswana (SMEC, 1991). There are many other examples of the use of the model in various consultancy reports that are not part of the scientific literature. Extensive use of the model has been made in South Africa since its development, culminating in the national water resource assessment exercise commonly known as WR90 where the model was regionalized throughout all the country’s so called quaternary basins (Midgley et al., 1994) which is now being updated (Bailey and Pitman, 2005). The Pitman model has found favour for water resource assessment, development and planning purposes in the region because of its relatively simple and flexible structure that can describe hydrological conditions in the region with some reasonable degree of confidence. The data demands can generally be met in a region that is haunted by problems of data scarcity. The most significant advantage is the rather coarse temporal scale at which the model operates. While data are scarce in the region, monthly records of evaporation, rainfall and runoff are not so difficult to get from various sources. Overall, the model simulations in the region have been considered acceptable by a wide group of scientists and practitioners. This has prompted a drive to explore the potential for full regionalization of the model following approaches already undertaken in South Africa (Hughes, 1997; 2004b), albeit with more robust parameter estimation procedures.

Besides the Pitman model there have been other models that have been developed and used in the region. The Variable Time Interval (VTI) model (Hughes and Sami, 1994) is one such model. It was also used quite extensively in basins of the region during the FRIEND project. Outside the FRIEND project applications, where it recorded mixed results, the VTI has only been applied in South Africa. The Namrom model, designed to work specifically in the basins of Namibia (de Bruine et al., 1993; Mostert et al., 1993), has not been tested in other basins of the region. Having identified aridity as an important climatic condition in the country the model was designed to simulate the hydrology of such areas including an allowance for varying, non-seasonal vegetation cover conditions and transmission losses to alluvial aquifers. Notwithstanding its limited range of applications, the model has been reasonably successful in Namibia. The model is quite simple, being a four parameter model based on a single equation of effective precipitation. With a sound conceptual basis the model has potential for wider use in similar conditions and other models could possibly benefit from its approach to the modelling of arid basins. Hughes and Meltzer (1998) added a dynamic vegetation cover to the Pitman model and achieved an improvement over the original model in Namibian basins.

The fully distributed, physically-based ACRU model, developed in South Africa at the University of KwaZulu-Natal (Schulze, 1986), has been applied mostly in the humid and temperate parts of the country. It is based on the idea of moisture accounting and uses multiple soil layers. Its application outside South Africa has been limited. The heavy data demands of the ACRU model would impact on its general use in the region in spite of the success it has enjoyed in the basins of South Africa where it has been used quite extensively.

Besides these “local” models, many other models, developed outside the region, have been used with varying success. Some of the most common of these are the HBV and the SHE models. The former was used in a number of selected basins in Zimbabwe during the Streamflow and Sediment Gauging and Modelling Project in Zimbabwe (GAMZ) with considerable success (SMHI, 2000). A daily Australian model, the Monash model (Porter and McMahon, 1971), was applied in Botswana for an integrated water development plan of southern Okavango (SMEC, 1987). However, models developed for local hydro-climatic conditions have tended to fare better.

3 THE PITMAN MODEL

3.1 Introduction

This chapter aims to provide a brief description of the Pitman model (Pitman, 1973) and its use in a database management and modelling framework system, developed at the Institute for Water Research, called SPATSIM (SPatial and Time Series Information Modelling, Hughes, 2002; Hughes and Forsyth, 2006). The model was borne out of the pioneer work of V.W. Pitman working in the Hydrological Research Unit at the Witwatersrand University. The development of the model was principally aimed at simulating "runoff in a form suitable for water resources appraisal" (Pitman, 1973; pp 1.7). The model is thus essentially a water resource assessment tool though some of its applications have often been deviated somewhat from the original plan for the model. Be that as it may, the model has acquitted itself well and has thus enjoyed relatively widespread use in the southern African region since the original version was developed in 1973. The Pitman model was originally designed as a conceptual lumped model but in more recent versions the model is semi-distributed. While the basic structure and form of the model has remained intact over the years, it has undergone a number of modifications. Two approaches have been evident with the later versions – the first being the use of nodes in order to better incorporate a broader spectrum of human influence in managed basins (Bailey and Pitman, 2005). The other route has been to use sub-basins in a distributed modelling approach, with the most recent version being the one in which explicit ground water routines have been added (Hughes, 2004a). This study was based on the latter version of the model. This chapter serves as an introduction to the model with the detailed structure and relevant algorithms being covered in Chapter 4.

3.2 The Pitman model

Figure 3.1 provides a flow diagram of the version of the Pitman model used in this study. The Pitman model is a monthly rainfall-runoff model whose inputs are monthly time series of rainfall totals and long term estimates of annual potential evapotranspiration. Though the model works on a monthly time scale the monthly rainfall totals are disaggregated into the four internal iterations over which the model works. The Pitman model is much like any typical conceptual model with tank type storages. Interception, soil moisture, and ground water are the three conceptual storages in the model. Rainfall first satisfies the interception storage before it finally reaches the surface as throughfall. The interception storage is decreased by evaporation at the potential rate. When the throughfall reaches an impermeable surface, direct overland runoff is generated; otherwise the water infiltrates the soil surface. If the surface infiltration capacity is less than the intensity at which the water is being supplied at the surface, the excess water runs off the surface in the Hortonian postulate.

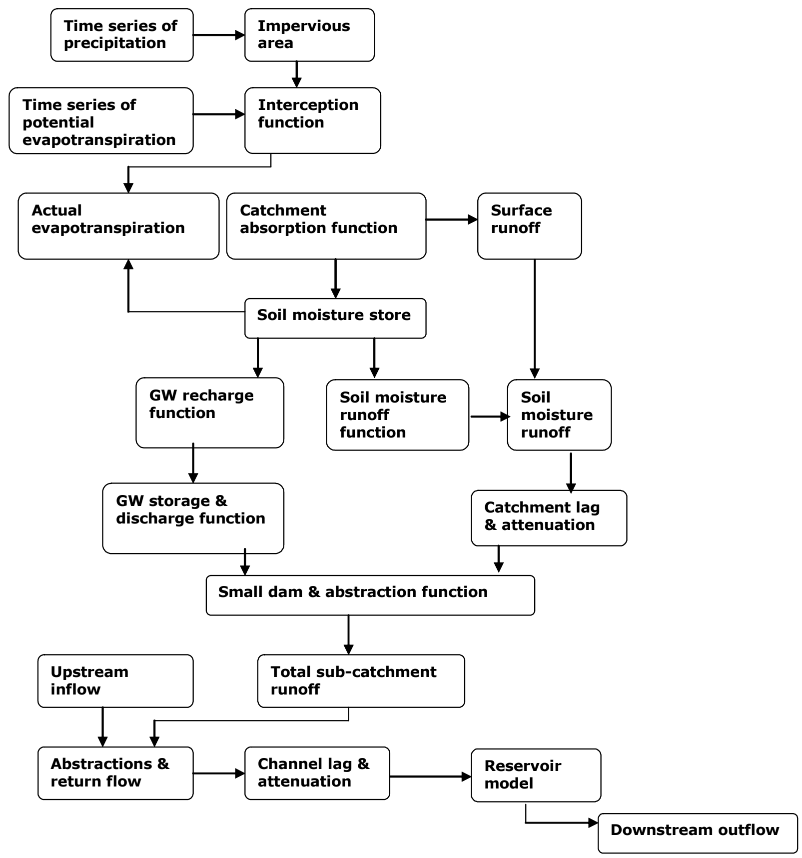


Figure 3.1 Flow diagram of the main components of the Pitman model (Hughes et al., 2006)

Infiltration is controlled by the soil surface conditions and is described by a triangular distribution with a minimum of ZMIN and a maximum of ZMAX that determines the proportion of the rainfall that can be absorbed by the basin surface, and therefore forms part of the soil moisture store, which can also contribute to runoff. The upper limit of the soil moisture storage is given by parameter ST. The moisture store is depleted through vertical drainage (i.e. recharge) through the zone of intermittent saturation via cracks/crevices/interstices, etc to the ground water store, lateral outflow to the channel and evapotranspiration directly from the soil and through vegetation.

Evapotranspiration from the moisture store is controlled by a parameter R, which describes the shape of the relationship between actual evapotranspiration losses and potential losses at different soil moisture levels. In simple terms this parameter determines the linear rate of decrease of the evaporative losses with decreasing levels of the moisture store (Hughes, 1994).

The amount of lateral runoff from the soil moisture store, with a maximum limit at saturation (ST) given by a parameter FT, is determined by a non-linear power function. Saturation excess overland flow is conceptualized in the model to occur when the soil moisture store is above saturation and this and the flow generated from the unsaturated store are routed together to the basin outlet by a simple lag function defined by a parameter TL. The recharge process increments the ground water store, while losses occur through the flow of the water to the channel as baseflow, flow to downstream sub-basins and evapotranspiration losses from riparian zones. The total runoff in the river channel can be subjected to transmission losses where the water will flow from the channel to the ground water store when there is a sufficiently high moisture gradient between the channel and the groundwater store. This is especially applicable to arid and semi-arid basins (Hughes, 2004a).

Since the model was designed for purposes of water resource assessment even in managed basins it also has functions that simulate the influence of man on the natural hydrology of a basin. There are routines to account for direct abstraction from the river itself and the ground water store for various purposes and provision is made for a proportion of the abstracted water to be returned to the river channel. Surface storage facilities are taken into account through the model's small dam routine (which affects runoff generated within a sub-basin) as well as a main reservoir water balance component that affects all runoff generated upstream. Both routines allow for evaporation losses and abstraction, while a limited number of abstraction and downstream release operating rules are built into the main reservoir component.

The current version of the Pitman model with ground water routines is quite heavily parameterized with a total of 41 parameters. The rationale is that the parameters "should be easier to evaluate for ungauged (or altered) situations because they are more meaningful in terms of real hydrological processes and can be related to measurable catchment characteristics" (Hughes, 2004b). However, most of the parameters can be estimated a priori from basin properties leaving some 11 free (calibration) parameters. The current study focuses on the development of estimation procedures for a subset of these calibration parameters. The next chapter considers all the parameters of the model in more detail with the relevant model algorithms and the current calibration procedures. It then develops a conceptual framework for their interpretation in physical terms that will be used for the quantification of the parameters. Table 3.1 gives a list of all the parameters of the model including some brief explanations.

Guidelines for the calibration of the parameters have evolved with the use (e.g. Middleton et al., 1981, Hughes et al, 2006) of the model from the initial parameter estimation guidelines given by Pitman (1973). In the water resources assessment study (Midgley et al., 1994) that included South Africa, Swaziland and Lesotho, regionalized parameter sets were developed for a total of 1946 so-called quaternary basins. These parameter values have provided pre-calibration initial estimates in the gauged basins and provide the best parameter value estimates for ungauged basins whose sizes are equal to the ones used to develop the regionalization (Hughes, 1997). The FRIEND project provided a platform for testing these calibration principles and the results were generally satisfactory, an indication that there is great potential for the regionalization of the model (Hughes, 1997). It is however pertinent to highlight that the model has so far been calibrated manually. There seems little incentive to change this approach given that there are strong reservations about the effectiveness of automatic calibration in achieving consistency in parameter values across a range of basins within a region where the accuracy and reliability of the input data are often dubious (Hughes, 2004b). However, the SPATSIM version of the model has a built-in facility that allows the automatic calibration of a number of parameters. This facility requires the specification of parameter limits that constrain the calibration process. Performance assessment in the model is through six objective functions, the Nash-Sutcliffe (1970) coefficient of efficiency (CE), the coefficient of determination (R^2) and the deviation of the mean of the simulated from that of the observed time series. These three are taken for both the untransformed and the natural logarithm transformed values to give a total of six performance measurement criteria.

While the original model was designed to produce simulated river runoff, the version of the model used in this study has additional outputs which include recharge, soil moisture, transmission losses, evapotranspiration, baseflow, interflow, surface runoff which can be used for both the assessment of model performance and as inputs to other models that are linked to the model in SPATSIM.

3.3 Use of the Pitman model in SPATSIM

This study will not endeavor a detailed description of the software and the reader is referred to publications by Hughes (2002), Hughes and Forsyth (2006) and Mwelwa (2004). The Spatial and Time Series Information Modelling (SPATSIM) software was developed at the Institute for Water Research (IWR) at Rhodes University as an improvement over its predecessor (HYMAS, Hughes et al., 1994) which lacked GIS functionality and was basically used for managing data for use with several different hydrological models. It is a database management and modelling framework specifically designed for hydrological and water resource system applications (Hughes, 2002; Hughes and Forsyth, 2006). SPATSIM uses some GIS functions and allows access to database tables for use with models through four data dictionaries. These allow the SPATSIM to be used as a data platform by different, even older, versions of models (Hughes and Forsyth, 2006). All spatial data loaded into the software through shapefiles whose associated attributes are stored in database tables. SPATSIM has a suite of internal facilities designed to allow the manipulation of data linked with the spatial elements. These facilities include routines for the import/export of data, addition/deletion of spatial features and/or attributes, data exchange protocols between SPATSIM users and a host of common hydrological data processing facilities. Examples of the last group include the generation of duration curves from time series and the generation of spatially averaged (over defined polygons) data using an inverse distance weighting method (Mwelwa, 2004). Besides these internal facilities SPATSIM also links with external models and data analysis programs that are individual entities developed outside the software. These include a generic time series data display and analysis program (called TSOFT, Hughes et al., 2000) and a collection of models of which the version of the Pitman model used in this study is one (Hughes and Forsyth, 2006).

Table 3.1 A list of the parameters of the Pitman model including those of the reservoir water balance model (Hughes et al., 2006).

Parameter	**Units**	**Parameter description**
RDF	Rainfall distribution factor Four	Controls the distribution of total monthly rainfall over
		model iterations
AI	Fraction	Impervious fraction of sub-basin
PI1 and PI2	mm	Interception storage for two vegetation types
AFOR	%	% area of sub-basin under vegetation type 2
FF Veg1		Ratio of potential evaporation rate for Veg2 relative to
PEVAP	mm	Annual sub-basin evaporation
ZMIN	mm month^{-1}	Minimum sub-basin absorption rate
ZAVE	mm month^{-1}	Mean sub-basin absorption rate
ZMAX	mm month^{-1}	Maximum sub-basin absorption rate
ST	mm	Maximum moisture storage capacity
SL	mm	Minimum moisture storage below which no GW
		recharge occurs
POW		Power of the moisture storage- runoff equation
FT	mm month^{-1}	Runoff from moisture storage at full capacity (ST)
GPOW		Power of the moisture storage-GW recharge equation
GW	mm month^{-1}	Maximum ground water recharge at full capacity, ST
R		Evaporation-moisture storage relationship parameter
TL	months	Lag of surface and soil moisture runoff
CL	months	Channel routing coefficient
DDENS		Drainage density
T	m^2 d^{-1}	Ground water transmissivity
S		Ground water storativity
GWSlope		Initial ground water gradient
AIRR	km^2	Irrigation area
IWR	Fraction	Irrigation water return flow fraction
EffRf	Fraction	Effective rainfall fraction
NIrrDmd	Ml yr^{-1}	Non-irrigation demand from the river
MAXDAM	Ml	Small dam storage capacity
DAREA	%	Percentage of sub-basin above dams
A, B		Parameters in non-linear dam area-volume relationship
IrrAreaDmd	km^2	Irrigation area from small dams
CAP	Mm^3	Reservoir capacity
DEAD	%	Dead storage
INIT	%	Initial storage
A, B		Parameters in non-linear dam area-volume relationship
RES 1–5	%	Reserve supply levels (percentage of full capacity)
ABS	Mm^3	Annual abstraction volume
COMP	Mm^3	Annual compensation flow volume

The first step in using the Pitman model in SPATSIM involves the preparation of all the relevant spatial coverages for the sub-basins, hydro-meteorological measuring networks and basin characteristics. Reference to the shapefiles can be included in a SPATSIM application, after which relevant attributes can be added and these populated with data for each of the spatial elements (points or polygons). The attributes can include single values (e.g. basin area), tables of data (e.g. parameter values, monthly evaporation) and time series (observed monthly flows, monthly rainfall data, etc). A model application is established by identifying the spatial elements to be included and associating appropriate SPATSIM attributes with the model input or output requirements. A TSOFT application can then be established to allow a graphical view of at least the observed and simulated time series data. The model results can then be evaluated through TSOFT statistically or graphically. The graphical evaluation methods include time series comparisons with a zoom facility, flow duration curve comparisons and scatter-graphs of observed and simulated flows. The goodness-of-fit between the simulated and the observed time series data may also be evaluated by way of visual inspection or by the calculation of at least six statistical objective functions (Fig. 3.2). If the results of the simulation are not satisfactory, the parameter values can be edited and the process repeated until acceptable correspondence between the observed and simulated flows is obtained.

An alternative version of the model allows the user to specify parameter value limits and step sizes for several as an additional input. The program then determines all possible parameter combinations, runs the model for each one and outputs a summary of the objective functions. This has been designed to allow a user to 'explore' parameter interdependences, look for optimal parameter value combination and address issues of equifinality and parameter identification in a relatively simple way.

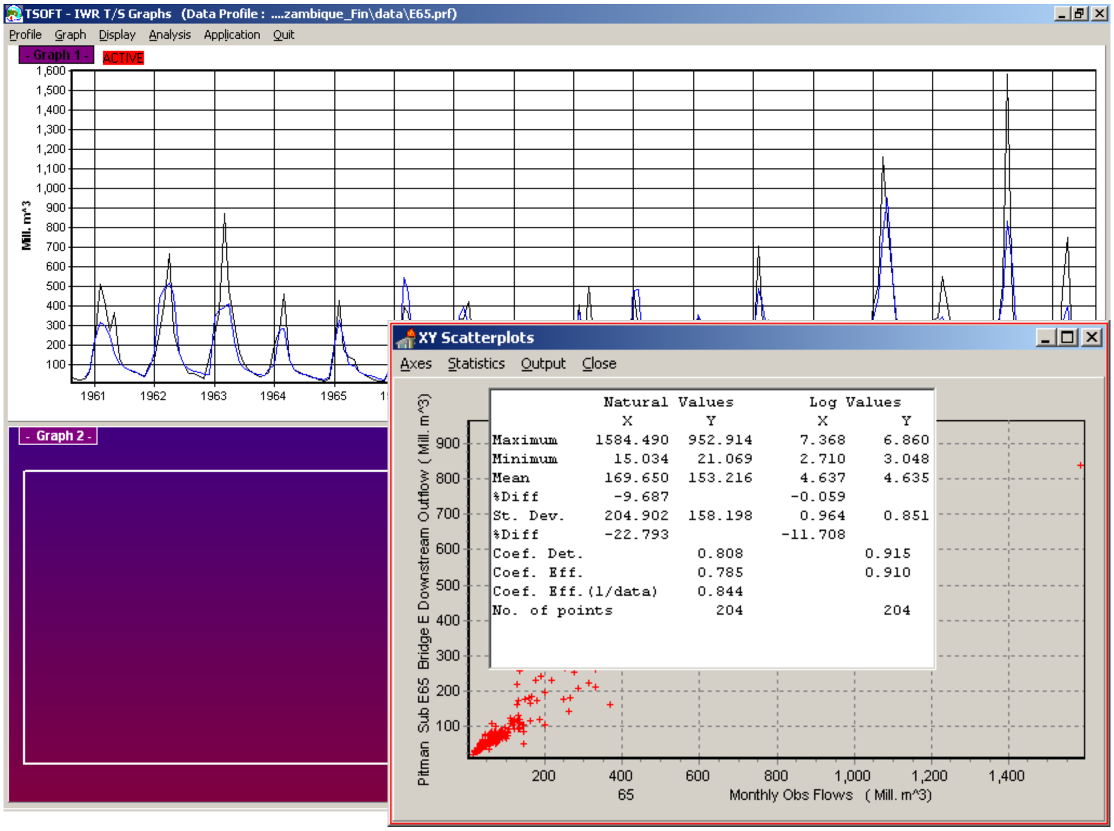


Figure 3.2 A graphical view of the observed (white) and the simulated (blue) in the top diagram and a statistical analysis of the goodness of fit in the TSOFT analysis program of SPATSIM.

4 PITMAN MODEL PARAMETER DESCRIPTIONS

4.1 Introduction

The purpose of this chapter is to provide a description of the Pitman model, through a detailed explanation of the parameters and associated model algorithms. The original model was developed in 1973 by Pitman (1973), but the version applied for the present study has undergone a number of modifications, the most substantial of which being the addition of revised procedures for simulating the interaction between surface and groundwater (Hughes, 2004a; Hughes and Parsons, 2005). For the purposes of this document the current version of the model is referred to as PITMGW. This chapter is structured by grouping parameters that have a similar purpose within the model and for each grouping the following information is provided:

i. the role of the parameters in the model and the associated model algorithms
ii. the physical interpretation of the parameter
iii. suggested principles for parameter quantification and calibration

The first set of sub-sections focuses on describing the parameters and algorithms and how they link to other components of the model. A great deal of the information contained within these first sub-sections is drawn from the original model manual (Pitman, 1973), as well as documents, published and unpublished, that have described subsequent modifications (Hughes, 1997, Hughes, 2004a; Hughes and Parsons, 2005 and others). The second set of sub-sections discusses any potential for physical interpretation of the parameters and therefore provides a conceptual basis for links to physical basin characteristics and properties. The third set of sub-sections refers to some of the principles that can be used for calibration of the parameters, but will also develop the theme of links to measurable physical basin properties and explore the available possibilities for quantifying parameter values from other information. However, before the individual parameter groups are discussed in detail, there are some general issues that need to be briefly highlighted.

4.1.1 Seasonal variations

The PITMGW model allows for the interception and infiltration parameters to have seasonal variations and these are controlled by setting July and January (specified as winter, or 'w', and summer, or 's', in the parameter descriptions) values for the parameter and including in a separate input stream the seasonal weights (W_m) using the following algorithm:

$$PAR_m = PARw + (PARs - PARw) * W_m \quad \text{.. 4.1}$$

where m represents the month subscript and PAR refers to the parameter being considered. Default values for the weights follow a sine curve shape with the value for $W_{July} = 0$ and $W_{Jan} = 1$. The weights can, of course, be modified by the user to produce almost any seasonal distribution shape. However, it should be noted that only one set of non-dimensional seasonal distribution weights is used.

4.1.2 Scale effects

One of the important issues to be able to recognize in the application of any hydrological model is whether or not any of the parameters are affected by basin scale effects. These effects may be part of the model formulation, or they may influence the way in which parameter values are quantified and their relationship with measurable basin properties. There may also be scale effects within the measured basin properties. While this topic will be discussed in more detail at a later stage in this document, it is important to recognize the potential for scale effects to impact on model results at an early stage. The Pitman model has generally been applied at scales of 10s of km^2 to basins over 10 000 km^2. Clearly, within that wide range of basin sizes there is a great potential for scale effects to play a role.

4.2 Rainfall Distribution Factor (RDF)

The model is designed to operate with input time series data with a monthly time resolution. In order to capture temporal variations in the rainfall input the model operates over four iterations (i.e. roughly 1- 7, 8 – 15, 16 – 23 and 24 – 30 days). The distribution of the total monthly rainfall is assumed to be controlled by a symmetric S-curve function that is dependent on the total rainfall and the rainfall distribution factor (RDF) parameter. This is a non-varying rainfall distribution parameter that determines how much rain is input in each iteration step. The parameter was introduced during the FRIEND project (Hughes, 1997). In the original version this value was fixed at 1.28, while lower values result in a more even distribution of rainfall, the effect being more pronounced for higher total rainfalls (Mwelwa, 2004).

4.2.1 Model structure

The model assumes relatively low rainfall in the first and last iterations and higher rainfall in the middle pair with the amount of rainfall being equal in the first and last iterations and also equal in the two middle iterations. The model distributes the monthly rainfall using the following equations:

$$y = x^n * [x^n + (1 - x)^n] \qquad 4.2$$

where y = cumulative rainfall / total rainfall; x = cumulative time / total time, and n is approximated by

$$n = RDF * (1.02 - W/P)^{-1.49} \qquad 4.3$$

where n = exponent related to the range of the maximum deviations above and below the uniform rate line, W, for cumulative daily rainfall for a given month;

P = total precipitation for the month, and $W = -2 + 1.3732 * (P + 1.6)^{0.8}$

The equations and numerical constants were derived empirically by Pitman (1973) from observations of daily rainfall distributions within South Africa.

4.2.2 Physical meaning of parameter

Clearly, any physical meaning of this parameter is associated with typical distributions of real rainfall during a month. This can be highly variable in some climate zones, depending on the type of rainfall events that occur and the number of rain days. Arid zones that experience infrequent events of relatively high intensity may be expected to require high values of RDF. Temperate zones with more frequent rainfall events might be expected to experience lower rainfall intensities and a more even distribution of rainfall within a month (lower values of RDF). Similarly, high rainfall totals during the wet season of sub-tropical zones might be expected to be relatively evenly distributed within a month. Mwelwa (2004) investigated the relationships between daily rainfall distributions and RDF parameter values for several rainfall stations in the Kafue basin, Zambia.

4.2.3 Calibration principles

There is a potentially strong interrelationship between the value of parameter RDF and the other model parameters that affect surface runoff. If the value of RDF is relatively high, the second and third iterations in the model will experience relatively high rainfalls which will generate greater volumes of surface runoff. A reduction in surface runoff during high rainfall months can therefore be achieved with a reduction in the value of parameter RDF or changes to the main surface runoff generation parameters (ZMIN, ZAVE and ZMAX). High rainfalls in the second and third iterations may also result in the maximum soil moisture storage level (ST) being exceeded and additional surface runoff being generated. RDF should not be considered as a calibration parameter and should ideally be established from a knowledge of temporal variations of rainfall within a month. The only reported analysis of daily rainfall distributions from representative rainfall stations within a basin for the purposes of estimating RDF can be found in Mwelwa (2004). She noted that a relationship seems to exist between monthly rainfall totals and the most appropriate value of RDF since there is a general increase in the evenness of the distribution of rainfall with an increase in monthly rainfall total (i.e. lower RDF value). While the model functions do allow for the distribution to become more even with higher rainfalls even with a fixed value of RDF, it appears that this may not always be sufficient. Eventually Mwelwa (2004) adopted a compromise RDF value that generally favoured the high rainfall months when the time distribution of the rainfall was expected to be more critical in terms of runoff generation. The present study will attempt to offer improved guidelines for establishing appropriate values for the RDF parameter within different regions of southern Africa.

4.3 Interception Parameters: PI1, PI2

A proportion of any precipitation input does not reach the basin surface because it is intercepted by the vegetation cover. The model has a routine to deal with this and the interception parameters are used to determine the proportion of precipitation that is lost through this process. The routine is based on the premise that:

i. the total rainfall on any rain day is concentrated in one storm event only, and
ii. all the intercepted rainfall is evaporated (at the potential rate, PEMAX) before the next rain day.

A single interception storage parameter, PI, was introduced in the original version of the model, while soon afterwards an allowance was made for a different interception storage over those parts of the basin covered by forest (parameter AFOR, the fraction of the basin under forest cover). During the FRIEND project (Hughes, 1997) the two values for interception storage were formalized in the model and the ability to define seasonal variations in one or more of the storages was introduced (see 4.1.1). While the main purpose of introducing the second storage parameter was to be able to simulate the impacts of afforestation, the two values can be used to represent any two dominant vegetation type groupings in a basin.

4.3.1 Model structure

Total monthly interception loss is assumed to be determined by interception storage capacity (PI) and the total rainfall. The following empirical equations are used within the model:

$$I = x * (1 - e^{yp}) \quad \ldots\ldots\ldots\ldots \quad 4.4$$

where I = total interception loss per month; p = total precipitation for the month and x, y are constants. For acceptable interception storage capacities (PI) varying from 0 to 8 mm, later measured in South Africa by Schulze (1995), x and y were approximated as:

$$x = 13.08 * PI^{1.14} \quad \ldots\ldots\ldots\ldots \quad 4.5$$

$$y = 0.00099 * PI^{0.75} - 0.011 \quad \ldots\ldots\ldots\ldots \quad 4.6$$

Equations 4.4 to 4.6 are used in the model algorithms to yield monthly total interception loss. The complete equation for this algorithm can be written as:

$$I = 13.08 * PI^{1.14} * [1 - \exp (p * (0.00099 * PI * 0.75 - 0.11))] \ldots\ldots\ldots \quad 4.7$$

It is important to recognize that interception loss becomes one component of the overall evaporation loss calculated by the model. However, while interception losses occur within the month of the rainfall and therefore affect the infiltration and surface runoff calculations, the other evaporative losses affect the soil moisture balance calculations and have a delayed effect. It can therefore be very important to get the correct balance between interception losses and 'real' evaporative losses.

4.3.2 Physical meaning of parameters

The process of interception is affected by the percentage of the ground covered by the vegetation and the leaf area index (LAI) of the vegetation type (Rutter et al., 1975). Both of these can depend upon the stage of development of the vegetal cover and the season of the year. It should also be noted that at the basin scale there will almost always be large spatial variations in interception capacity. There are a number of literature sources that have documented interception losses for different vegetation types (for example, Rutter et al., 1975; Schulze, 1995; Valente et al., 1997; Zeng et al. 2000; Hall, 2003). A direct comparison between the parameter values and measured interception capacity is somewhat confused by the model assumption that the stored water evaporates completely in a single day. In reality, within a monthly time step model, the extent to which this assumption can be considered valid will depend upon the typical patterns and distribution of rainfall within a month. If the total monthly rain falls in concentrated periods of several days it is likely that the model will over-estimate interception losses.

4.3.3 Parameter estimation principles

Traditionally the PI parameters have not been calibrated and in South Africa it has been the practice to use a value of 1.5mm for the natural vegetation condition and 4 mm for that proportion of the basin that is under plantation forest. Unfortunately, this approach does not take into account the substantial variations in vegetation cover and Leaf Area Index (LAI) that occur across the sub-continent. In addition, a value of 8mm generates monthly interception losses under some rainfall regimes that are excessive. As more information on vegetation cover characteristics becomes available from satellite data, for example (DeFries, et al., 1999) it should be possible to determine improved methods of estimating the interception parameters directly. As noted in the previous sub-section (4.3.2) it may also be necessary to account for regional differences in within-month rainfall characteristics. A Markov model of the probability of a rain day and a daily interception model can be used to estimate interception loss (de Groen, 2002; de Groen and Savenije, 2006). The parameters of the Markov model could be taken from hydrologically similar areas where studies were undertaken (see Hughes et al., 2006). The use of this method deserves further exploration since the model currently uses default values of 1.5 mm for natural forest and 4 mm for reforested areas.

4.4 Infiltration Parameters: AI, ZMIN, ZAVE, ZMAX

The model assumes two components of surface runoff generation:

i. Precipitation falling on an impervious surface adjacent to a stream into which the surface discharges directly. This is calculated as the product of the monthly rainfall and a parameter (AI) representing the proportion of the basin that is impermeable.
ii. From infiltration excess surface runoff.

To estimate the second, within a moderate to large size basin it is necessary to recognize that moisture absorption is likely to be spatially variable and depends on vegetation, soils and geology (Pitman, 1973). The infiltration parameters describe the absorption capacity of the basin in response to different rates of rainfall input. They are rates used to determine the proportion of the monthly rainfall input that is absorbed by the basin and therefore determine the amount of surface runoff generated for any given rainfall input. The model makes use of a triangular distribution of basin absorption rates varying from a minimum value of ZMIN to a maximum value of ZMAX. In the original model, the distribution was assumed to be symmetrical (ZAVE = (ZMIN + ZMAX) / 2), however, more recent versions allow for a non-symmetrical distribution by introducing ZAVE as a parameter (Fig. 4.1). In the PITMGW model version ZMIN is allowed to vary seasonally (see Section 4.1.1). Rainfall totals below ZMIN do not generate runoff and all moisture is absorbed, while higher rainfalls will progressively generate higher runoff. Rainfall rates greater than ZMAX will have that portion of the distribution above ZMAX all contributing to runoff generation.

4.4.1 Model structure

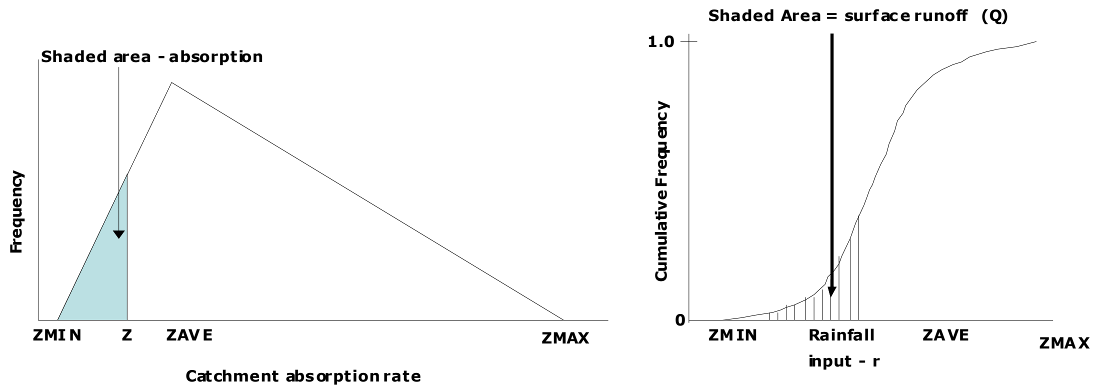


Figure 4.1 Illustration of a left-skewed non-symmetrical triangular frequency distribution of basin absorption rate, Z (left side) and the cumulative frequency curve illustrating the proportion contributing to surface runoff generation (right side).

The parameters are used to estimate basin absorption in the following manner based on the properties of a triangular distribution of the absorption rates. For any given absorption rate, Z mm per month

For ZMIN ≤ Z ≤ ZAVE,

$$\text{absorption} = (Z - ZMIN)^2 / (ZMAX - ZMIN)(ZAVE - ZMIN) \qquad 4.8$$

and for ZAVE < Z ≤ ZMAX,

$$\text{absorption} = 1 - [(ZMAX - Z)^2 / (ZMAX - ZMIN)(ZMAX - ZAVE)] \qquad 4.9$$

Absorption will be zero for Z ≤ ZMIN and will be equal to one (which is the maximum since the total area under the triangle is unity) when Z = ZMAX. These equations are then used to generate the excess of precipitation over infiltration which becomes surface runoff, Q.

Given any rainfall input rate, r

for ZMIN ≤ r ≤ ZAVE

$$Q = (r - ZMIN)^3 / 3(ZMAX - ZMIN)(ZAVE - ZMIN) \qquad 4.10$$

for ZAVE ≤ r ≤ ZMAX

$$Q = r - ZAVE - A + \frac{(ZAVE - ZMIN)^3/}{3(ZMAX - ZMIN)(ZAVE - ZMIN)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ 4.11$$

where $A = \dfrac{ZMAX^2 (r - ZAVE) + ZMAX * (ZAVE^2 - r^2) + 1/3\,(r^3 - ZAVE^3)}{(ZMAX - ZMIN)(ZMAX - ZAVE)}$

and for r › ZMAX,

$$Q = r - ZAVE - A + \frac{(ZAVE - ZMIN)^3}{3(ZMAX - ZMIN)(ZAVE - ZMIN)} \quad \ldots\ldots\ldots\ldots\ldots\ 4.12$$

with $A = \dfrac{ZMAX * ZAVE^2 - ZMAX^2 * ZAVE + 1/3\,(ZMAX^3 - ZAVE^3)}{(ZMAX - ZMIN)(ZMAX - ZAVE)}$

It should be noted that the values of the infiltration parameters are closely linked to the rainfall distribution factor (RDF – Section 4.2) that controls the way in which the total monthly rainfall is distributed over the four model iterations. Lower values of RDF will reduce the rainfall rate in the two main wet periods, while increasing it in the other two periods. Within a complete month the relationships between generated runoff, the RDF parameter and the infiltration parameters can be quite complex.

4.4.2 Physical meaning of parameters

The infiltration parameters represent the spatially integrated process of infiltration. During any rainstorm event for a given basin, the rate of infiltration generally decreases from a high rate to a minimum steady rate (infiltration capacity) which should approximate the soil's saturated hydraulic conductivity and is dependent upon a range of soil properties such as structure, porosity, texture, macro-pore density, surface sealing, etc). The whole process of infiltration at the basin scale is highly complex and strongly influenced by surface and sub-surface basin characteristics. It is therefore extremely difficult to ascribe direct physical meaning to the parameters. However, it is possible to suggest some guidelines:

- Basins with large spatial variations in soil properties will be expected to have relatively large differences between ZMIN and ZMAX.

- Coarser textured and well drained soils (sands) are expected to have higher values for all parameters than finer textured soils (clays and loams).
- If surface sealing is an important process at the beginning of storm events, ZMIN might be expected to have a low value (i.e. small rainfall amounts generate some runoff).
- Arid basins with thin soils and low values of soil moisture storage are expected to have relatively low infiltration parameter values, which reflect the fact that the infiltration capacities of thin and stony soils will be quite low.

Despite the observations above, it is not a simple task to assign physical meaning to the infiltration parameters of a monthly model that is designed to operate at relatively large spatial scales. Although there are many estimates of infiltration capacities for different soil types within the literature (e.g. Warrick and Amoozegar-Fard, 1979; Brakensiek, 1977), it is a different matter to assess the relationships between these and the values of ZMIN, ZAVE and ZMAX.

4.4.3 Parameter estimation principles

The previous sub-section suggests that it will be difficult to determine any direct estimation methods for the infiltration (or basin absorption) parameters. At the same time, these parameters are of critical importance in semi-arid and arid areas where sub-surface runoff generation processes are considered to have a very small influence. Pitman (1973) provides some calibration guidelines based on comparisons of observed and simulated runoff characteristics. In general terms, an increase in ZMIN results in a decrease in simulated mean annual runoff (MAR) and an increase in standard deviation. It will also affect the seasonal flow distribution in that periods of higher rainfall will generate greater runoff than the low rainfall months. An increase in ZMAX has no effect on runoff reliability or seasonal flow distribution but will decrease MAR (Pitman, 1973). Calibrating the value of ZAVE is possibly even more difficult and establishing a suitable value will depend on the degree of asymmetry assumed in the spatial variation of basin absorption rates.

There is clearly a need for further investigations into the methods and sources of information that could be used to estimate the values of the infiltration parameters. Unfortunately, many soil maps do not provide a great deal of hydrologically relevant information and are often more related to agricultural potential. Many of the ACRU model (Schulze, 1986) soil parameters are based on the South African land type maps which combine topographic position, soil type and soil depth into a classification system. There is therefore potential for the same information (at least in South Africa) to be used for the infiltration parameters of the Pitman model. In other areas of southern Africa it may be possible to make use of the FAO soil maps (FAO, 2003) combined with additional information on topography and vegetation.

4.5 Soil Moisture Storage Parameters: ST, SL

The moisture storage content of a soil increases due to infiltration. ST represents the maximum value of this storage expressed in mm. If ST is exceeded in any time step, the balance of the precipitation contributes to runoff. Decreases in the level of this storage result from evaporation losses (section 4.6), runoff (section 4.7) and recharge to ground water (section 4.8). Pitman (1973) introduced this parameter as one that determined the ability of basins to regulate the runoff from a given precipitation input. Though not explicitly stated by Pitman (1973), parameter ST in the original model must include some allowance for groundwater storage. This is because both the soil moisture and groundwater runoff volumes generated by the original model (using parameters FT, GW and POW) were extracted from the same available storage ($SL < S \leq ST$). While ST has been maintained in all versions of the model through to PITMGW (Hughes, 2004a), the addition of an explicit groundwater store as a separate component and the redefinition of GW implies that the original meaning of ST (and therefore possibly its value) has been altered from the original version. In the original model SL represented the lowest moisture storage level at which all runoff ceased. In the PITMGW version of the model SL is only used to limit groundwater recharge and there is no lower limit used for runoff generation (see sections 4.7 and 4.8). Part of the motivation for this change was that SL was almost always set to zero in conventional applications of the model and thus had little significance (Hughes, 2004a).

4.5.1 Model structure

In the original model version the soil moisture runoff at ST was FT and at lower moisture levels (S) the runoff was determined using a non-linear relationship as illustrated in Figure 4.2. If runoff was less than a parameter GW (groundwater component of runoff) all the outflow was lagged slowly using parameter GL, while the runoff proportion greater than GW was lagged using a different lag parameter (TL) in order to generate a more rapid response (see section 4.9). In PITMGW the redefinition of SL and GW has meant that all soil moisture runoff is lagged using the same parameter (TL). GW no longer plays a role in soil moisture runoff generation and has been re-defined as groundwater recharge. Fig 4.2 also illustrates the relationships in the revised PITMGW model (Hughes, 2004a).

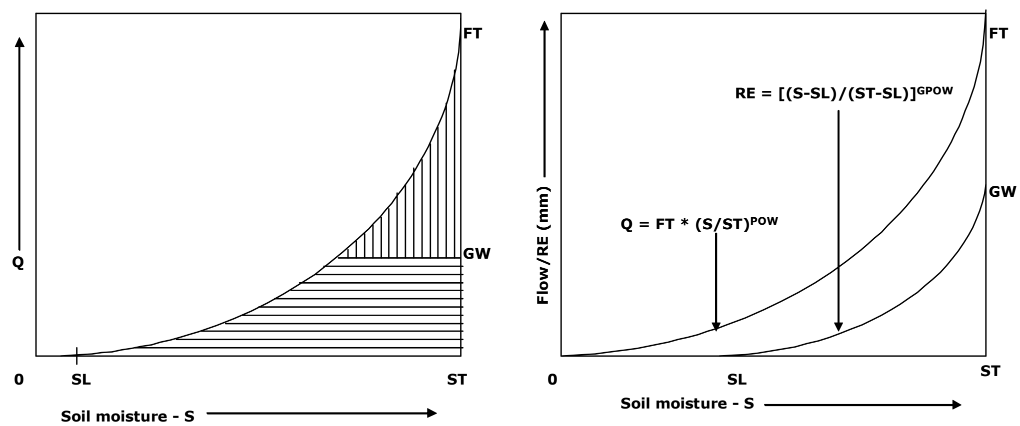


Figure 4.2 Illustration of the relationship between ST, GW and FT (on the left) and the subsurface runoff generation parameters as used in PITMGW (SL, ST, FT, POW and GPOW).

In summary, ST is used by the soil moisture component of the model to regulate:

i. evapotranspiration from the soil moisture storage
ii. soil moisture runoff
iii. groundwater recharge

4.5.2 Physical meaning of parameters

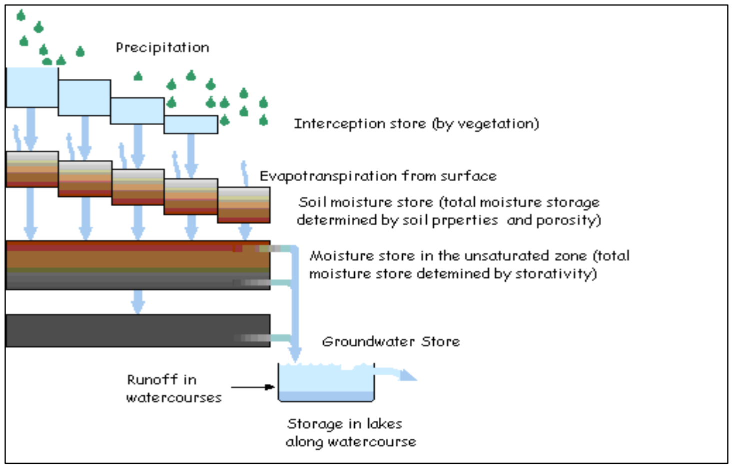


Figure 4.3 Illustration of the conceptualization of the moisture storages of the PITMGW model

Figure 4.3 illustrates the conceptualization of the moisture store and the movements of water into and out of the store. One of the important aspects associated with the physical interpretation of the parameters is to identify exactly what this storage represents. In the first instance it clearly represents moisture stored in the soil. The only other drainage component in the model refers to outflows from the saturation zone below the water table. This implies that ST must also represent the storage within the unsaturated zone below the soil and above the water table. In some basins within southern Africa it has been suggested that this zone may play an important role in runoff generation. As large parts of southern Africa are underlain by fractured rock systems, the ST parameter must account for the moisture storage potential of these fractures. It is quite possible for a large proportion of the fractures above the water table to be saturated despite the rock matrix material being unsaturated. Within the soil the maximum amount of moisture that can be held (moisture holding capacity) is largely determined by the soil's porosity. This in turn will depend on the depth, texture and structure of the various soil layers (Rawls et al., 1982; Cosby et al. 1984) Soils developed from geological formations that are easily weathered would be expected to have deeper soils and higher values of ST, as would areas where the prevailing climate promotes deep weathering.

Within the unsaturated zone, below the soil, fractured formations have the capacity to store moisture within cracks and fissures. Storativity refers to the volume of water that a permeable geological formation will absorb or expel from storage per unit surface area per unit change in head and is equal to the product of specific storage and aquifer thickness. The storativity of a fractured formation will be made up of the available storage in the fractures as well as that in the rock matrix. In the context of the PITMGW model, it is the storativity of the fractures that has the most relevance to the value of ST. This is because it is assumed that it will be that component of storage in the unsaturated zone that could potentially contribute to runoff (see further discussion in section 4.8). This assumption is based on the premise that it is possible for some part of the fracture system to be saturated while the formation as a whole is unsaturated. This could lead to relatively rapid interflow, which is combined with soil moisture runoff within the model. The model variable S (with a maximum of ST) therefore represents a combination of the soil moisture storage (ST_{soil}) as well as the storage in the fracture zone (ST_{unsat}) that has the potential to contribute to interflow. It is therefore possible to suggest an approach to estimating ST from a knowledge of the soils and unsaturated zone physical properties. For example, in principle at least, given a soil with depth of 1.5m and porosity of 40%, the maximum amount of moisture stored could be estimated by:

$$ST_{soil} = 0.4 * 1500 = 600 \text{ mm} \quad \ldots\ldots \quad 4.13$$

If the depth of the unsaturated zone below the soil zone is 20m with a storativity representing the fractures of 0.001 then the maximum moisture storage could be estimated by:

$$ST_{unsat} = 0.001 * 20000 = 20 \text{ mm} \quad \ldots\ldots \quad 4.14$$

with ST = ST_{soil} + ST_{unsat} = 600 + 20 = 620mm.

This type of approach will be discussed in greater detail in the next chapter.

4.5.3 Calibration Principles

One of the critical issues associated with calibrating the value of ST is related to high rainfall months when there is the possibility that ST will be exceeded in any single model iteration. The depth of effective rainfall that exceeds the available storage becomes runoff and if the value of ST is too low, this runoff volume can become excessive. Conversely if the value of ST is too high the volume of runoff generated may be too little as ST may never be reached. This is especially so in the drought season of most southern African basins. Clearly there will be a high degree of interrelationship between, *inter alia*, the value of the rainfall distribution parameter, the surface runoff parameters and the value of ST. In arid areas, where soil moisture runoff can be assumed to be negligible, the value of ST will largely determine the frequency with which the moisture store will be filled and runoff generated through the overflow process. In the revised PITMGW version of the model ST will also affect patterns of groundwater recharge, which may influence channel losses in arid areas. In wetter areas with sustained baseflow, the value of ST will influence the variability of soil moisture runoff, low values producing a higher degree of variability than higher values.

The previous section proposed one method by which the value of ST could be estimated from information on physical basin properties and this approach has been evaluated as part of this research project (see later sections in this document). However, even if successful in some basins, it will be difficult to apply in areas where the required information on soils and the unsaturated zone is limited.

4.6 Soil Moisture Runoff Parameters: FT, POW

FT refers to the runoff generated from the soil when the moisture level (S) is at its maximum value (ST). The relationship between generated runoff and moisture level is illustrated in Figure 4.2. As already noted, SL was defined as the lower limit of soil moisture at which outflow ceases, but is not used within this function in the revised model. Runoff from the soil moisture store is assumed to be regulated through a non-linear relationship between discharge and soil moisture. This runoff was originally divided into components, each of which was lagged differently in the routing component of the model. In PITMGW with the revision of the GW and SL parameters this is no longer the case and all runoff from soil moisture is lagged using parameter TL. This relationship is assumed to be adequately represented by a simple power (defined by parameter POW) function.

4.6.1 Model structure

Refer to Fig 4.2 for illustration. Runoff from the soil moisture store is determined by the following equation;

$$Q = FT*(S/ST)^{POW} \qquad \text{4.15}$$

where S is the current soil moisture store, and Q = FT if POW = 0 or S = ST

As the soil moisture store is increased by continued infiltration during a rainstorm event the amount of interflow also increases to the value FT when the moisture store is at its maximum (ST). Beyond ST all excess rainfall is converted directly to runoff. In reality therefore FT cannot be greater than ST as this would mean runoff greater than the amount of moisture available to generate that runoff. All other factors being constant an increase of POW will result in an increase in discharge.

4.6.2 Physical meaning of parameters

It is necessary to recognize that the soil moisture runoff function represents the total volume of drainage (interflow) from the zone above the water table at the basin scale. Section 4.5 referred to the physical meaning of ST and the fact that it can represent both soil moisture and unsaturated zone storage. FT must therefore represent the maximum possible runoff from these two sources. At small scales the total runoff that can drain from a soil will be approximately the difference between porosity and field capacity, which will obviously be dependent on soil texture and structure and other soil properties. However, at the basin scale many other factors play important roles and the complexities of topography and spatial variations in soil type and depth will affect local drainage, as well as opportunities for ponding and re-infiltration. In a similar way, the generation of interflow in the unsaturated zone will be dependent upon, among other things, vertical variations in permeability, fracture orientation, as well as the degree of interconnectivity and connectivity with the surface channel network. When the total basin is less than saturated (S < ST) there will be areas (especially those in the valley nearer to the channel) that remain above field capacity (due to drainage from upslope, for instance), while other areas will be drier. POW can therefore be assumed to represent the relationship between total basin moisture status and the spatial distribution of this moisture. The concepts are similar to the probability distributed principle of Moore (1985) which has been used in several lumped or semi-distributed models including the VTI model of Hughes and Sami (1994). The relationship will be dependent upon topography and the spatial arrangement of soil types and depths and is further complicated in the Pitman model by the long time scale of modelling. While the information that is likely to be required to establish direct parameter estimation procedures (for FT and POW) is unlikely to be generally available, this issue is considered to be worth further investigation and will be re-visited later in the document.

4.6.3 Calibration principles

These parameters should not effect runoff generation in arid and semi arid areas where sustained baseflow does not exist. While soil moisture runoff may exist at small scales within such basins, re-infiltration and evaporation loss processes preclude the generation of runoff at the basin scale. Parameter FT is therefore normally set to zero and runoff generation largely controlled through the surface runoff function involving parameters ZMIN, ZAVE and ZMAX. However, in areas of intermittent or sustained baseflow, FT and POW assume great importance in the model calibration process. Establishing suitable values can be a complicated task, largely due to the interdependence with many of the other parameters. In many respects, the addition of the groundwater interaction routines in the PITMGW version of the model has complicated the calibration of FT and POW even further. There are now two functions that can generate sustained baseflow and model users need to ensure that they understand the type of runoff response that will result from changes to the two sets of parameters. In practice, it would seem advisable to calibrate the parameters of the more slowly responding groundwater functions against late dry season or drought low flows and then focus on FT and POW to obtain the best fits for the recession at the end of the wet season. However, these principles require further investigation and will be revisited later in the document. If the PITMGW version of the model is being applied in a basin with an existing calibration and parameter set for an earlier version of the model, it is important to recognize that FT no longer includes a component of groundwater outflow. The obvious assumption would be that this parameter should therefore decrease in value, but no guidelines are currently available to indicate by how much.

4.7 Groundwater Recharge Parameters: SL, GW, GPOW

SL was previously defined as the minimum storage below which no soil moisture runoff occurs, but was conventionally set to zero for most applications. Its importance in the model was therefore naturally irrelevant. In PITMGW SL has been redefined as the lower limit of soil moisture below which no groundwater recharge occurs. The original meaning of GW was referred to in section 4.5 and has been redefined in the PITMGW version to refer to the upper limit of the groundwater recharge rate (in mm per month) at moisture state ST. To quantify recharge at different moisture levels (S) a new parameter (GPOW) was introduced to define the form of the relationship between recharge and current moisture storage (Figure 4.2).

4.7.1 Model structure

Groundwater recharge is computed in the model according to the relationship:

$$RE = GW * [(S - SL)/(ST - SL)]^{GPOW} \quad \ldots\ldots 4.16$$

where RE is the monthly recharge rate in mm, S is the current soil moisture storage level in mm and SL is the lower limit of soil moisture state of the soil below which no groundwater recharge occurs.

4.7.2 Physical meaning of parameters

GPOW describes the shape of the relationship between moisture stored in the unsaturated zone and the volume of recharge. It is therefore very similar to POW and can be expected to reflect similar physical relationships. As the total moisture status of the basin as a whole declines, the proportion of the basin with soil moisture states above field capacity will decrease. At the small scale field capacity is assumed to be the lower moisture content limit for vertical drainage. At some moisture level it is possible that there will be no parts of the basin that have soil moisture states above field capacity and no fractures in the unsaturated zone containing sufficient water to generate vertical drainage. This level of basin storage would be equivalent to SL, the point at which no recharge will be generated. The maximum recharge rate should be linked to the same factors that affect the maximum soil moisture runoff rate (FT), including soil texture and structure. However, while topography will play a major role in the determination of FT (slope gradients in areas with low topography will be insufficient to generate much lateral drainage), it will play a lesser role in the vertical recharge process.

The information typically available to define the vertical structure of the unsaturated zone and its relationship with surface topography is rarely detailed. While it is therefore possible to identify the physical relevance of these runoff generation parameters, it will not normally be possible to use such concepts in quantifying parameter values.

4.7.3 Calibration principles

The value of SL, as was the case in the earlier version, can normally be set to zero without compromising the results. The rationale being that the rates of recharge at low soil moisture are small and have little influence on the total water balance of the basin (Hughes and Parsons, 2005). The previous sub-section indicated that while GW and GPOW have conceptual physical meanings, direct estimates of their values will almost always be difficult. The recharge process is understood to be highly non-linear as well as spatially variable (Scanlon et al., 2002). Toth (1963) clearly explains the impact of topography on local and regional flow paths and recharge is thus generally expected to occur in topographic highs and discharge in topographic lows in the more humid environments. In the arid alluvial valley zones recharge is assumed to be focused in topographic lows such as channels of ephemeral streams (Scanlon et al., 2002). Hydro-geomorphic regions (Meyboom, 1967) can easily be delineated using GIS and digital elevation models (DEM) and could be called upon to assist in the identification of areas of active recharge and discharge (Hatton, 1998). While the identification of active recharge areas is important, it is the quantification of the recharge rates needed for the model that will always be difficult. In the absence of relationships between recharge rates and basin properties, it should be possible to use annual or monthly recharge estimates against which the values of GW and GPOW could be calibrated. Bredenkamp et al., (1995), Baron et al. (1998) and Xu and Beekman (2003) give recharge values for South African basins based on a number of different assessment methods. The model would be run and GW continuously adjusted until the recharge result equals or approximates the values given in the literature. Unfortunately, the literature rarely contains information on annual variations or seasonal distributions of recharge, both of which could be very different for similar annual means.

4.8 Evaporation Parameters: R, AFOR, FF

The evaporation function depends on the current month's potential evaporation value relative to the month with the highest potential evaporation together with the values of parameters R and FF. R defines the relationship between the ratio of actual evaporation to potential evaporation and the level of the soil moisture store (S). R essentially determines the shape of a linear relationship assumed between actual losses and potential losses at different moisture storage levels and its meaning has been maintained through all versions of the model. The value of R varies between 0 and 1 inclusive (see Fig. 4.4). A further parameter (FF) has been introduced in several of the more recent versions of the model and represents an evaporation scaling factor for a second vegetation type, frequently used to represent plantation forestry. The proportion of the basin area covered by the second vegetation type is given by the parameter AFOR.

4.8.1 Model structure

Figure 4.4 illustrates the relationships between basin evapotranspiration (E) and soil moisture storage level (S) for the limiting conditions of R=0 and R=1. The full equation for the evapotranspiration function is:

$$E = PE*[1 - \{1 - R*(1 - PE/PEMAX)\}^{-1} * (1 - S/ST)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 4.17$$

Including the effect of the second vegetation type the total evapotranspiration E_{Total} is given by:

$$E_{Total} = E * FF * AFOR + E * (1 - AFOR) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 4.18$$

Figure 4.4A illustrates that when R=0 evapotranspiration continues even when the soil moisture content is very low, regardless of the potential evaporation for the month. In contrast, when R=1 evapotranspiration ceases at higher values of S as the potential demand decreases (Fig. 4.4B).

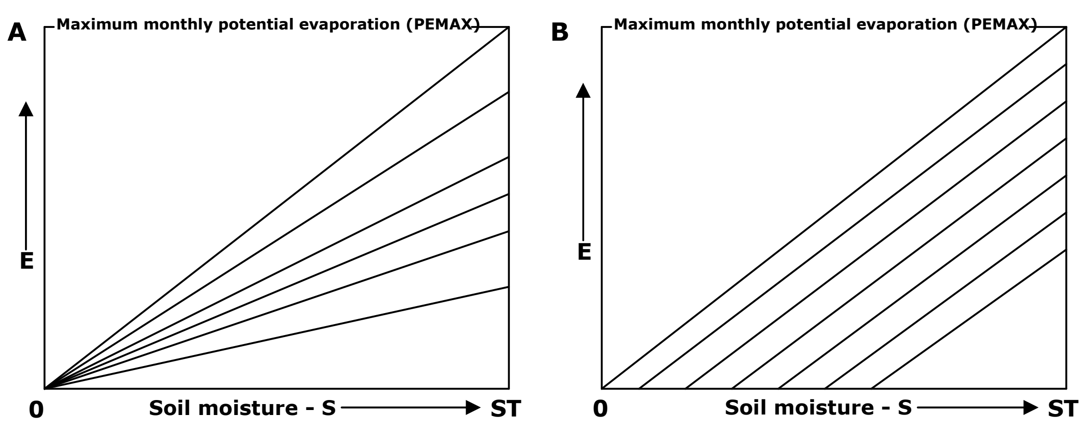


Figure 4.4 Relationship between basin evaporation (E) and soil moisture (S) for R= 0 (A) and R = 1 (B)

4.8.2 Physical meaning of parameters

Though the parameters are not physically-based, loose connections can be made between their values and rooting density and depth, as well as the way in which soil texture affects soil water tension at various levels of moisture storage. The fact that ST may also represent moisture stored in deeper unsaturated zones (section 4.5) should also be taken into account. R is therefore a parameter that should reflect the effectiveness of vegetation to extract water from the moisture store. FF is a simple scaling factor that allows the second vegetation type to have greater evapotranspiration losses than the remainder of the basin area.

4.8.3 Calibration principles

The calibration guidelines provided by Pitman (1973) for the parameter are based on the understanding that an increase in R will increase simulated runoff and result in a more uniform seasonal flow distribution. The value of the parameter should be influenced to a great extent by the type and density of vegetation. Rooting depth and density will determine the rate of depletion of the soil moisture. A potential source of confusion that may, however, arise as a consequence of the conceptualization of the ST parameter should be addressed. In situations where the deeper unsaturated zone dominates the moisture storage capacity of basin over the soil zone (e.g. in areas of thin soil cover such as some arid basins) the depletion of moisture by evapotranspiration may not be fast enough and therefore interflow will remain high. This is possible given that evapotranspiration is assumed to be ineffective in the deeper unsaturated zone especially beyond the rooting depth. Increasing the parameter value would not affect the amount of subsurface flow.

4.9 Runoff Routing Parameters: TL, GL, CL

The parameter TL refers to the runoff time lag in months that is applicable to the surface and soil moisture runoff components. GL was formerly the time lag that was applied to the groundwater component of runoff in the original model and was assumed to be always greater than TL. In the PITMGW version parameter GL is no longer used and the revised groundwater functions (see section 4.10) act as a routing reservoir. CL is a parameter that has been added to perform channel routing in large basins, where even at the monthly time scale delays and attenuation may occur as runoff is routed from upstream through downstream sub-basins.

4.9.1 Model structure

The lag parameters are used within the Muskingum routing equation in which the weighting factor (x) is set to zero to represent reservoir type storage attenuation. The normal Muskingum equation is given as follows:

$$0_2 - 0_1 = C_1 (I_1 - 0_1) + C_2 (I_2 - I_1) \quad \text{4.19}$$

where $C_1 = \Delta t / [K (1 - x) + 0.5\Delta t]$ 4.20

$$C_2 = (0.5\Delta t - Kx) / [K (1 - x) + 0.5\Delta t] \quad \text{4.21}$$

with subscripts 1 and 2 referring to previous and current months' runoffs respectively and the other variables assuming the following meanings for the purposes of the model; 0 = monthly runoff total at basin outlet, I = instantaneous monthly runoff, Δt = routing period, K = lag of runoff, X = weighting factor and for x = 0, $C_1 = \Delta t / (K + 0.5\Delta t)$ and $C_2 = 0.5\ C_1$.

TL therefore represents the K value appropriate to routing runoff generated within a specific sub-basin, while CL represents the value appropriate to routing upstream runoff through a downstream sub-basin.

4.9.2 Physical meaning of parameters

It is reasonable to assume that the value of TL will be related to the sub-basin size and response rate, which in turn will be related to topography (sub-basin slope), drainage density, dominant type of runoff (surface or soil moisture/interflow). However, at the monthly time scale it is difficult to identify any clear physical associations.

CL will be mainly related to the size and channel length of sub-basins. However, it could also be related to channel and riparian characteristics such as slope, in-stream vegetation, floodplain width, etc. As with TL it is difficult to identify any clear physical relationships based on generally available data.

4.9.3 Calibration principles

Higher values of TL will reduce peak flows, generate slower recessions at the end of the wet season and sustain baseflows during the dry season. Previous recommendations (Midgley et al., 1994) have suggested that a value for TL of 0.25 can be used in most situations. However, it is not clear whether this approach is applicable to sub-basins that are either very small (< $50km^2$) or very large (> 100 $000km^2$). TL is therefore normally not calibrated and there seems little justification to change this approach. CL is only applicable to large basins and there is very little existing experience of its use.

4.10 Groundwater Accounting Parameters: DDENS, T, S, RWL, GWSlope, RipFactor

These parameters are an integral part the groundwater discharge component of the PITGW version of the model. DDENS refers to the drainage density of the basin and is expressed as a ratio of the total channel length to the basin area given in km km^{-2}. The assumption in the model is that the drainage density includes only those channels that are likely to receive groundwater discharge. This would exclude many tributary channels that only receive surface runoff or interflow. T refers to the transmissivity (m^2 d^{-1}) of the aquifer and is a product of permeability and saturated aquifer thickness, while S refers to the storativity, a measure of the capacity of the aquifer to store water.

The RipFactor is the riparian strip factor parameter. This defines the volume of water loss through evaporation close to the channel margin. This is achieved through evaporation loss from the channel bed and banks and through evapotranspiration of near-surface ground water by riparian vegetation. In the model the RipFactor is given as a percentage of the total slope element width over which the evapotranspiration process is active. The Rest Water Level (RWL) parameter represents the maximum depth below the channel that the aquifer is assumed to reach. At this level all ground water movement is assumed to cease. Its conceptual definition is related to the ground water geometry calculations in the model. The GWSlope parameter represents the regional ground water gradient that is used to determine drainage from a sub-basin to downstream sub-basins.

4.10.1 Model structure

In combination with the basin area, the DDENS parameter is used to define the geometrical representation of the ground water. The DDENS parameter is used to determine the number of channel and slope elements that will be used in the estimation of ground water outflow (see Fig 4.5)

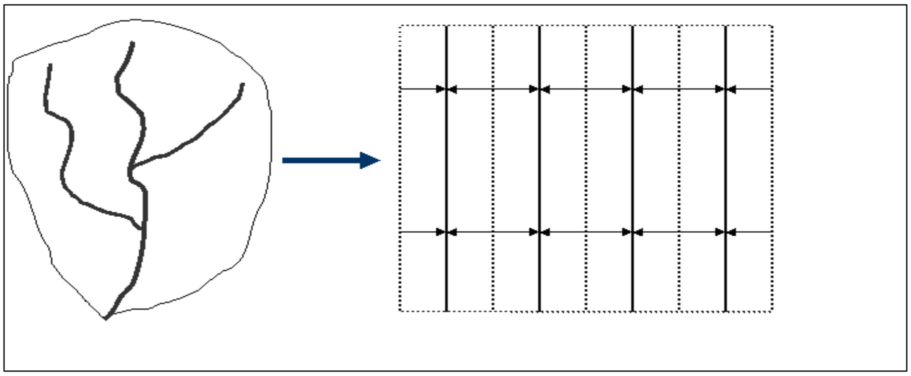


Figure 4.5 Conceptual representation of drainage in the basin where the channels are of unit length and DDENS of 4/sqrt (Area) (Solid lines are channels, dashed lines are drainage divides and the arrows show drainage directions). Adapted from Hughes (2004a).

The total length of channel (TCL) expected to receive ground water discharge is initially calculated from:

TCL = sqrt (Area *0.5 / DDENS)………………........…………....................... 4.22

The number of contributing slope elements (NSlope) is initially estimated from:

NSlope = 2 * Area * DDENS / TCL ….............….........................………... 4.23

A check is made that the NSlope variable is at least 2 and then NSlope is corrected to the nearest even valued integer, after which the total width of each slope element is calculated from:

Width = Area / (TCL * NSlope) ……………...............................……............ 4.24

The total width of each drainage slope is divided into 'near channel' (40% of the total width) and 'remote from channel' (60% of width) compartments which are modeled separately (see later). The main reason for including two compartments within each slope element is related to the way in which ground water abstractions are expected to impact on discharge to the channel (see later section). Increments to, and losses from the ground water aquifer are used in the volumetric calculations based on the slope element widths and lengths coupled with simulated lateral (i.e. across the slope elements) ground water gradients and the storativity parameter. As the volume of water changes within the conceptual aquifer, the two gradients (near and remote from the channel) will also change. The gradient variables are used to determine movement of water within the aquifer (from the remote from channel compartment to the near channel compartment), as well as discharge from the aquifer to the channel. In the near channel compartment the aquifer is assumed to be always in contact with the river channel and therefore situations where the water table is below the river are simulated with negative gradients (Fig. 4.8). The volume of water in an aquifer compartment is calculated from simple geometry as:

$$\text{volume} = \frac{(\text{drainage width})^2 * \text{drainage length} * \text{gradient} * \text{storativity}}{2} \quad 4.25$$

where the drainage width and gradient variables refer to the values in either the near channel or remote from channel compartments. Within the model, increments to the aquifer occur as recharge from the surface component of the model, ground water drainage (see later section) from an upstream sub-basin and flow from the channel if the near channel compartment gradient is negative. The first two of these are added to the two compartments in proportion to their widths (40:60). Losses from ground water occur as drainage to downstream sub-basins, evaporation losses in the riparian strip (near channel only), discharge to the channel (near channel only), discharge to the near channel compartment from the remote from channel compartment and abstractions.

The model constrains one end of the near channel compartment water level to be at the channel, while the end of the remote from channel water level must be at the same point as the other end of the near channel water level (see Fig. 4.6). It is therefore possible for several conditions to exist within the geometry of the conceptual aquifer:

i. Positive gradients within both compartments. Under this condition the remote from channel compartment drains to the near channel compartment, the near channel compartment drains to the channel (and is subject to riparian evaporation losses) and both compartments drain to downstream sub-basins at a rate determined by a regional ground water gradient parameter (GWSlope). No channel losses are possible due to the existence of a positive near channel gradient (Fig 4.6, diagram A).
ii. Positive gradient in the near channel compartment, negative gradient in the remote compartment. This situation would only exist where abstractions from the remote compartment have drawn the water level down, or a combination of recharge and channel losses have increased the volume (and therefore gradient) in the near channel compartment. Downstream drainage (to the next sub-basin) would still occur from both compartments unless the point joining the two compartments is at the RWL. Note that the near compartment does not discharge to the remote compartment (Fig 4.6, diagram B).
iii. Positive gradient in the remote from channel compartment and negative gradient in the near channel compartment. Drainage from the remote from channel component to the near channel compartment will occur, as well as riparian evaporation losses (while the point joining the two compartments is above the RWL) from the near channel. Channel losses to the near channel compartment can occur (Fig 4.6, diagram C).
iv. Negative gradients in both compartments. No internal drainage will occur, while the two compartments will operate as before under negative gradient conditions (Fig 4.6, diagram D).

The initial gradient for both compartments is taken as the regional ground water gradient (GWSlope). The model is run through the complete time series once with this starting value and then the model is re-run using the compartment gradients at the end of the first run as new starting values. This approach has been adopted to avoid having to specify starting values and seems to generate stable results in most cases. Within a slope element compartment, the discharge (Q) in m^3 (to the next compartment or to the channel) is calculated by:

Q = Transmissivity * Gradient * Time Step * Length........................ 4.26

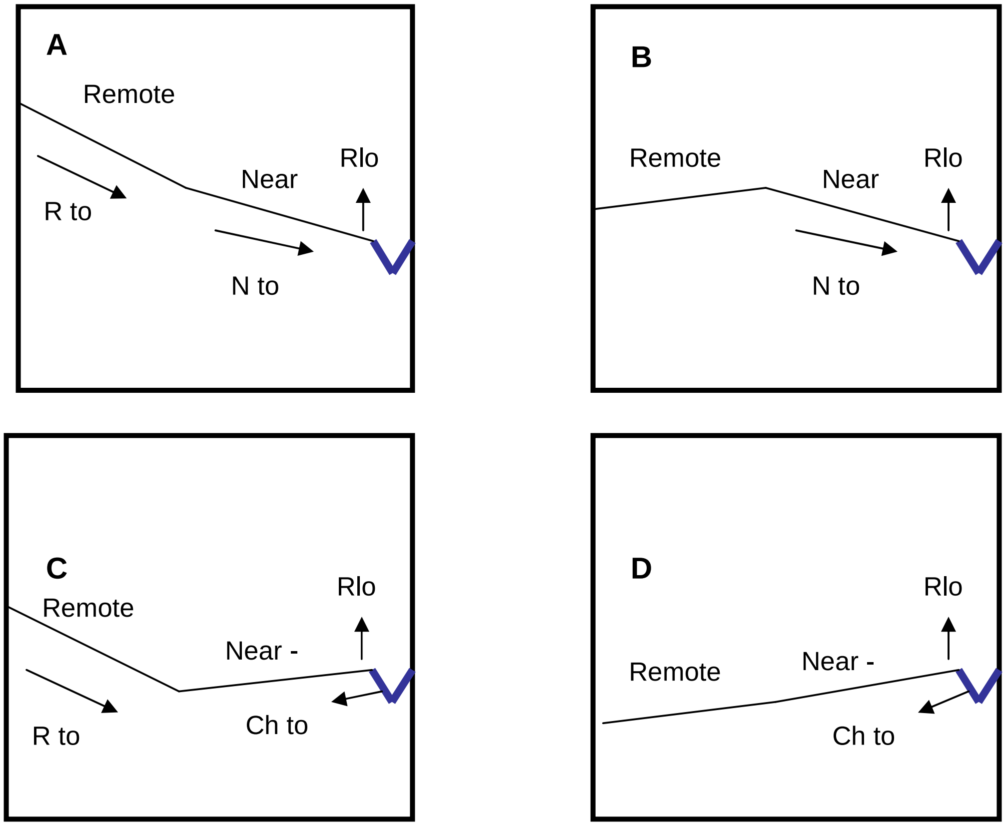


Figure 4.6 Illustration of the channel width compartments and the different conditions that can exist within the geometry of the conceptual aquifer (R is remote from channel compartment, N is near channel compartment, Ch denotes the channel and Rlos refers to evaporation loss from riparian vegetation. The arrows indicate the direction of movement of water)

The length is the same for both compartments and is equivalent to the channel length adjacent to each slope element (i.e. total channel length * 2 / NSlope). Riparian evapotranspiration from the near channel compartment is based on losses from an area representing the proportion of the total slope element given by:

Loss area = Slope width * Channel length * RipFactor/100................ 4.27

and

Evapotranspiration losses = Net Evapotranspiration * loss area 4.28

Net evapotranspiration is assumed to be potential evapotranspiration less rainfall (negative values are assumed to be zero). The evapotranspiration losses are first taken away from any calculation of discharge to the channel, while the remainder is taken from the volume in the near channel ground water compartment. If the near channel compartment gradient is negative a reduction factor is calculated based on the current gradient compared to the gradient at RWL($grad_{RWL}$):

Reduction factor = ($grad_{RWL}$ – current gradient)/ $grad_{RWL}$ 4.29

The reduction factor is used to reduce the evapotranspiration losses such that riparian losses decrease as the ground water gradient becomes increasingly negative. Discharge to downstream sub-basins is based on the following equation and is removed proportionally from the two compartments:

outflow = Transmissivity * Regional Gradient * Time * Slope Width....4.30

Under negative near channel compartment gradient conditions, the reduction factor (given above) is used to reduce the rate of downstream outflow. In summary the modelling process for each model iteration step is as follows:

i. The recharge is calculated and the associated volume of water added to the near and remote compartments, taking into account the storativity.
ii. The gradients during the previous time step are used to estimate outflow from the remote compartment to the near compartment, the outflow from the near compartment to the channel and the regional ground water gradient used to calculate the outflow to the downstream basin. The riparian evapotranspiration losses are calculated, as are any channel transmission loss inputs to ground water and any abstraction losses from ground water.
iii. The new volumes of water in the two slope compartments are calculated and used to calculate the gradients for the next time step. All of the volumetric water balance calculations are interpreted into simple geometry calculations to determine the gradients.

4.10.2 Physical meaning of parameters

It should be recognized that the approach used is a compromise between representing the real processes of sub-surface flow and using simple geometry to represent the aquifer. The spatial scale of modelling should also be taken into account when considering the modelling approach, as well as the physical interpretation of the parameters. Most of the aquifers in southern Africa occur within fractured rock systems with very low primary permeability and large spatial variations in ground water characteristics. On the other hand, a great deal of the quantitative information available is from isolated observations obtained during borehole drilling operations. All of the ground water parameters have direct physical meaning. However, the way in which they are quantified may depend upon the particular circumstances in any one region or basin.

4.10.3 Calibration principles

While it will be frequently necessary to calibrate some, or all, of the ground water parameters, it is recommended that the initial values be established on the basis of the best available hydrogeological information. The drainage density parameter determines how many slope elements are included in the definition of the aquifer geometry, the total length of channel receiving ground water, as well as the width of the drainage slope elements. Lower drainage densities clearly result in shorter channel lengths and therefore less ground water discharge per month for the same values of other parameters. Lower DDENS values also result in smaller total sub-basin outflow widths and therefore lower rates of ground water drainage to downstream sub-basins. An initial value of 0.4 is deemed ideal for headwater basins of ill-defined geometry. Lower DDENS values (0.3 or 0.2) would be ideal in situations of elongated basin shape where the transmissivity values are relatively high so that excessive drainage to downstream basins does not occur.

Under normal circumstances the storativity and transmissivity parameters should be quantified on the basis of the rock type and its degree of fracturing. Values for these parameters are documented in standard ground water texts (e.g. Xu and Beekman, 2003) and can frequently be used with little adjustment. Under certain circumstances it may be necessary for the transmissivity values to reflect the rates of water movement in fracture zones rather than in the aquifer as a whole.

The RipFactor parameter has less direct physical meaning than most of the other ground water parameters. It should reflect the areal extent and type of riparian vegetation which is likely to either use near-surface ground water directly or intercept ground water discharge contributions to streamflow.
The rest water level is mainly relevant to semi-arid basins where the ground water table is consistently below the channel bed. It will have very little effect on the overall model results in most cases, but could impact on the extent to which large abstractions (relative to mean annual recharge) from ground water can be maintained.

The regional ground water slope only affects the ground water drainage to downstream sub-basins, which is typically a minor component of the water balance of southern African basins. However, this process could be locally important and under such circumstances it is essential to quantify this parameter realistically.

For South African basins, estimates for most of the ground water parameters are available from a database, Groundwater Resource Assessment II (GRAII), developed under a Department of Water Affairs and Forestry project (DWAF, 2005)

4.11 Channel Loss Parameter: TLGMax

It is well understood that streamflow can be lost from the channel to the aquifer under circumstances when the level of ground water near the channel falls below the level of the channel. However, the satisfactory quantification of this process has eluded many hydrologists working in semi-arid basins and there are few guidelines in the literature on the best approach to use to establish suitable model algorithms. It is important to note that there are two potential channel loss processes. The first is channel losses from the incremental runoff generated within a sub-basin, while the second is channel losses generated within the main channel passing through a sub-basin which also affects upstream flows from other sub-basins. In this model they are treated separately but using a similar algorithm. The only model parameter affecting channel loss is TLGMax, which refers to the maximum runoff loss from the whole sub-basin in mm month^{-1}. For the main channel losses (affecting upstream flow) this is re-interpreted as a maximum channel loss in million m^3 month^{-1}.

4.11.1 Model structure

The model calculates two components of channel loss in downstream basin which receives inflows from an upstream basin. First is the channel loss from the runoff being generated within the sub-basin being modeled (incremental runoff). The second component is the channel loss from flow in the main channel.

To calculate the channel losses to incremental runoff the following scheme is used in the model. Three other variables are needed and these are MAXQ, TLQ and TLG. MAXQ is the maximum runoff (in mm) for the sub-basin being modelled and is automatically estimated by the model during the first run. The variable is set to a default value of 20mm at the start of the first run. TLQ is a variable estimated from the current month's runoff (Q) and its value is calculated using the following equations (Fig. 4.7):

For Q/MAXQ < 0.25

$$TLQ = 0.5 * [\tanh(2.5*(Q/MAXQ - 0.25)) + 1.0] \quad \text{.......................... } 4.31$$

and if Q/MAXQ ≥ 0.25

$$TLQ = 0.5 * [\tanh(6*(Q/MAXQ - 0.625)) + 1.0] \quad \text{........................... } 4.32$$

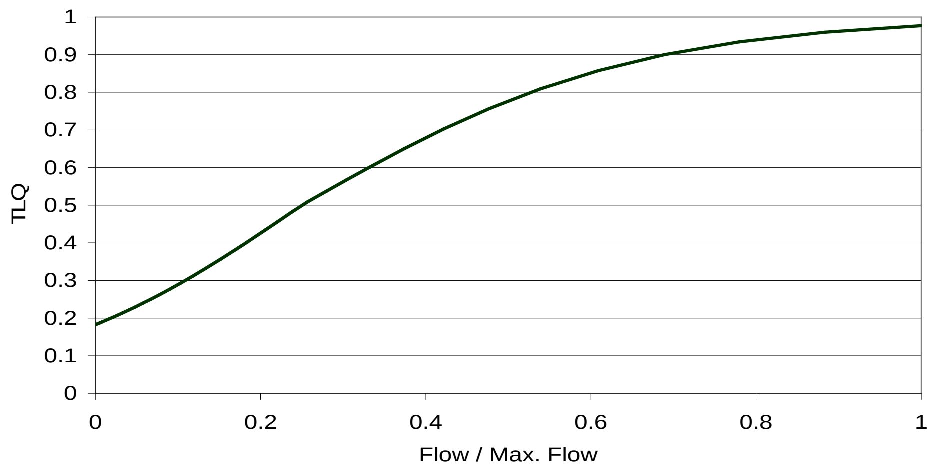


Figure 4.7 Shape of the power relationship between current month discharge (mm), relative to a maximum value (20mm in this case) and a model variable, TLQ.

TLG refers to the current gradient (Grad) relative to a maximum that is defined by 70% of the gradient at RWL (RWLGrad). It is thus a measure of the head difference between the channel and the groundwater (i.e. groundwater gradient of the near channel slope element) and they are related to each by a power function (Fig. 4.8).

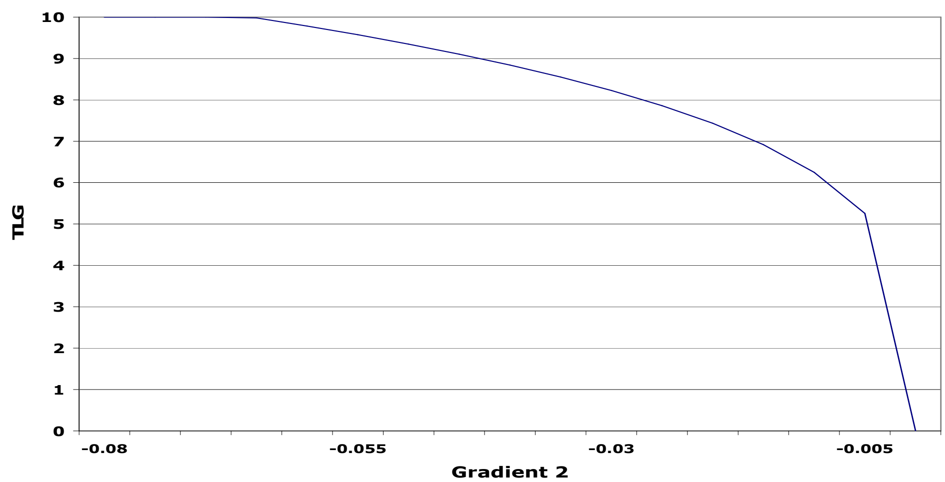


Figure 4.8 Shape of the power relationship between current down slope gradient and a model variable, TLG. The maximum value of TLG is defined by a model parameter.

TLG is estimated as follows: If Grad < 0.7* RWLGrad, then TLG = TLGMax, otherwise

$$TLG = TLGMax * [grad / (0.7*RWLGrad)]^{0.25} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 4.33$$

With these two variables the drainage from the channel runoff within the sub-basin can be estimated. This is the product of the TLQ * TLG. This volume of water is then removed from any available runoff within the sub-basin and added to the lower slope element, the reasoning being that any channel losses would obviously be a gain in the slope adjacent to the channel which is the lower slope element. The maximum channel loss will occur when the gradient of the lower slope element is at 70% of the gradient at RWL and when the sub-basin runoff is at its maximum value.

To estimate the channel loss to upstream runoff passing through the sub-basin (cumulative inflow channel losses) the same functions as described above are assumed to hold. However this time they are applied to the upstream inflows to the sub-basin. The groundwater gradient component (TLG) is retained though TLGMax now represents a maximum volume of channel loss from upstream inflow (in Mm^3) and, for convenience sake, will be denoted TLGMax_Inflow. This is calculated using the TLGMax parameter as:

$$TLGMax\_Inflow = TLGMax * (MAXQ\_Inflow / MAXQ) \quad 4.34$$

where MAXQ_Inflow is the maximum upstream inflow. Like MAXQ, MAXQ_Inflow is set to a default value of 20mm multiplied by cumulative upstream basin area with subsequent recalculation for the second run from the data simulated during the first run.

For the calculation of TLQ the same equations are used with MAXQ replaced by MAXQ_Inflow and Q representing upstream inflow in any one month. The cumulative inflow channel losses are estimated at the start of a single month's simulation and are then subtracted from the upstream inflow. This subtracted volume is then added to the lower slope element groundwater storage in equal amounts over the model iteration steps.

4.11.2 Physical meaning of parameter

It is difficult to ascribe any real physical meaning to this parameter and the only possible interpretation would be through observed maximum channel losses using a network of gauges.

4.11.3 Calibration principles

The estimation of this parameter value will never be simple, largely because of the highly non-linear nature of the channel loss process. The estimated channel losses in the model are also dependent upon the current months flow rate and near channel ground water compartment gradient. The parameter is clearly irrelevant in wetter basins where the ground water gradients are always positive. The only possibility for calibration would occur where nested gauged basins exist and where the downstream basin contributes little in terms of incremental runoff. However, even in such circumstances it has been found to be difficult to establish a pattern that can be satisfactorily simulated by the model algorithm. With respect to the incremental basin losses, the loss function can be used in situations where runoff is known to be generated in headwater parts of the basin, but frequently does not survive to the basin outlet. This allows the model user the flexibility of generating runoff internally, but losing some of the runoff before it continues to downstream sub-basins.

4.12 NON-NATURAL PARAMETERS

The model simulates the influence of various water resources developments on the natural streamflow and the parameters used in the model for these routines will be referred to as 'non-natural parameters' for convenience. There are two subgroups relating to water use and impoundments (reservoirs) on the streams. These parameters were not part of the original version of the model but are additions during subsequent versions. The parameters that help simulate the impact of impoundments on the river are MaxDam, DAREA, A and B and those relating to water use are Airr, IWR, NIrrDmd, IrrAreaDm and EffRf.

4.12.1 Water Use Parameters: Airr, IWR, IrrAreaDmd, NIrrDmd, EffRf

Besides flow reduction by transmission loss and evaporation processes which are simulated by the model, direct river abstractions for agricultural, domestic and industrial purposes are common. The PITMGW model has routines for differentiating direct abstractions from the river for irrigation and non-irrigation purposes.

Airr refers to the run–of-river irrigated area in a sub basin given in km^2. This governs the potential demand on river runoff by determining the size of area that can be irrigated. It is used in conjunction with a model attribute which describes the monthly distribution of irrigation depth (in mm) required. The effective rainfall parameter (EffRf) reduces the irrigation depth requirement by this proportion of the rainfall occurring within a month. IrrAreaDmd refers to the total area irrigated from small dams and is expressed in km^2, whose water demand is wholly satisfied from the small dams within the sub-basin. The monthly distribution of requirements is the same as for the run-of-river irrigation. NIrrDmd refers to the annual volume of non irrigation demand given in million litres (Ml). The non irrigation parameter is based on a specified annual demand value and is used together with a model attribute that fixes the monthly distribution of demand. The model attribute is made up of 12 rows (pertaining to months) and 4 columns (Monthly Distribution Weights, Monthly Irrigation Demand (mm), Monthly Water Demand (fraction) and Ground Water Demand (fraction)). The first column of data is used to distribute seasonally different parameter values for all the months, the second to determine the depth of monthly irrigation water demand, the third to distribute the annual non-irrigation water use value and the fourth to distribute the annual groundwater abstractions. IWR refers to irrigation water return flow. The rationale is that a certain proportion of the water abstracted for irrigation will return to the river system in any given month. The parameter is expressed as a fraction. Naturally this parameter would include both estimated water loss through seepage and the actual measured return flow in some large irrigation schemes where structures are available for measurement. EffRf is the effective rainfall fraction which is the proportion of a rain input that goes directly to satisfy irrigation demand i.e. that proportion useful for crop production. It therefore is important in determining the level of water demand on the river for irrigation purposes. If there is more moisture obtained from direct rainfall then the model simulates less irrigation demand from river flow.

4.12.1.1 Model structure

Direct abstraction from river runoff takes place to satisfy the irrigation demand of the irrigated area (Airr) when the amount of rainfall is not sufficient to cover its irrigation water requirements. The model treats this abstraction as a loss (Qloss) from the streamflow (given as a volume). The following algorithms are used to calculate Qloss. Firstly, an irrigation deficit (Irr Def) is established using EffRf and total monthly depth of irrigation demand (Irrm) (determined using a model attribute as given in the section above) as follows;

$$\text{Irr Def} = \text{Irrm} - \text{rain} * \text{EffRf} \quad \text{.......... 4.35}$$

The potential volume of runoff loss from the river to irrigation ($Qloss_{pot}$) is given by;

$$Qloss_{pot} = \text{Airr} * \text{Irr Def} \quad \text{.......... 4.36}$$

The actual loss from river runoff ($Qloss_{act}$) will vary depending on the streamflow levels. Thus, if streamflow ≤ $Qloss_{pot}$, then

$$Qloss_{act} = \text{streamflow} \quad \text{.......... 4.37}$$

i.e. all the available flow is taken away for irrigation and the model simulates zero outflow, otherwise $Qloss_{act} = Qloss_{pot}$ and

$$\text{outflow} = \text{Streamflow} - Qloss_{pot} \quad \text{.......... 4.38}$$

In reality however, a component of the abstracted water finds its way back to the river system by seepage or as a result of the irrigation scheme design which allows return flow. The model simulates this as a seepage volume as follows;

$$Seepage = Qloss_{act} * IWR \qquad 4.39$$

The net effect in the model is an increase of the streamflow and a reduction of $Qloss_{act}$ by an amount equivalent to the seepage value.

After taking care of the irrigation demand the model then simulates the non irrigation demand from the remaining streamflow (outflow above). The monthly distribution of non irrigation demand (MNIrrDmd) is estimated from the total annual demand (NIrrDmd) using a monthly distribution model attribute for the sub basin (mdist) as given;

$$MNIrrDmd = mdist * NIrrDmd \qquad 4.40$$

This volume is abstracted from river runoff (outflow from irrigation demand calculations) using the following argument. If MNIrrDmd ≥ outflow, all the flow is taken to satisfy it and the total loss from the stream ($Qloss_{tot}$) is given by;

$$Qloss_{tot} = Qloss_{act} + outflow \qquad 4.41$$

The model then simulates zero flow in the river. If MNIrrDmd ‹ outflow, then

$$Qloss_{tot} = Qloss_{act} + MNIrrDmd \qquad 4.42$$

and river flow is simulated as the difference between outflow and MNIrrDmd.

4.12.1.2 Calibration principles

Some water management authorities issue water use certificates/rights/permits to registered users for a specific abstraction level per given time. Such information may include both groundwater and river flow abstractions. This information is useful in estimating the amount of water used in any given sub-basin. It should be noted however that such information may give an indication of water allocation in the basin only but not actual abstractions. While such information can be available for large water use schemes like large scale commercial irrigation there is a dearth of information from small scale subsistence users. However the fraction used by this group is usually small and could safely be ignored or reasonably estimated. There is therefore normally no need to calibrate these parameters as they will be available from national or regional water management or agriculture departments. The volume of water abstracted for domestic water use in rural areas is even more difficult to measure accurately but this could be estimated by indirect methods (Wallingford, 2003).

The other data required are those that relate to the area under irrigation. While reasonably accurate estimates of the area under irrigation in a basin can be found, there still remains the problem of quantifying the volume of water used. While such programs as CROPWAT (Smith, 1992; Allen et al, 1998) could be used, not all irrigators follow the strict guidelines as given by the program and many irrigate on an ad hoc basis to supplement the natural rainfall in order to obtain yields in excess of the norm for dry-land crops (Midgley et al, 1994). However, in practice this information is not always available and some assumptions have to be made about the parameter values.

The effective rainfall and return flow parameters are the most difficult to estimate. Effective rainfall is influenced by a range of factors including rainfall intensity characteristics, soil properties and management practices. Of the many methods available to estimate EffRf, the US Department of Agriculture Soil Conservation Service method (Wallingford, 2003) is the most popular and can be used given the relevant rainfall and evapotranspiration data.

The measured surface component of IWR is not a problem. However the seepage component is difficult to estimate. IWR may be estimated from the type of crops, stage of development, soil properties, irrigation efficiency and evapotranspiration. None of these are easy to estimate. It is probably worth exploring the possibility of using subsurface flow (interflow) to estimate the seepage component. This would be based on the premise that the water returned to the river system would go via this process. The factors that influence surface flow would also impact on the return flow process.

The data on groundwater abstractions is easily available where the records for borehole drilling are available or where a database for groundwater use is available.

4.12.2 Reservoir Parameters: DAREA, MAXDAM, A, B

DAREA relates to the proportion of the sub basin commanded by the small dams. It is given as a percentage (%). The runoff generated in this area is assumed to initially satisfy the available reservoir storage before being able to contribute to flow at the outlet of the sub basin. Water held in the dams is subject to loss through

(i) abstraction for irrigation purposes which is thus controlled by the parameter IrrAreaDmd and a model attribute of monthly distribution of depth of irrigation water demand, and

(ii) evaporation which is controlled by a non linear relationship between area and volume and the monthly potential evaporation demand.

MaxDam refers to the capacity of the small dams' storage. This is the volume that needs to be satisfied before the basin area commanded by the reservoir starts to contribute to outflow from the sub basin. A and B are the parameters in the non linear capacity-area relationship of the reservoir. The capacity and area of a reservoir are assumed to be related in the following manner:

$$Area = A * Vol^{B} \quad \text{.......... 4.43}$$

where Area is given in m^2 and volume (Vol) m^3.

4.12.2.1 Model structure

The dam storage is incremented by all sources of runoff generated within a sub-basin. The model simulates these contributions as losses from both groundwater and surface flow. The routine starts off by estimating the dam's storage potential (St Potential) as the flow depth (in mm) required for the dam to fill to capacity. The modeling sequence is such that surface runoff generated in the upstream area of the dam is used first to fill up the storage. If the dam fills up from the surface runoff then the balance is allowed to flow over the dam and contribute to flow at the basin outlet Otherwise the all the flow is absorbed by the dam and baseflow generated in the area is used to fill up the dam. If the dam then fills up, the balance of the baseflow will contribute to flow at the basin outlet. If the dam fails to fill up, then all the baseflow is absorbed by the dam storage and the area will not contribute to the flow observed at the basin outlet.

The processes described above determine the capacity of the dam at the beginning of any month. This capacity (damvol) is reduced due to evaporation and abstraction. The evaporation occurs from the dam's surface area at the potential rate and the volume lost to this process is determined by the parameter R. The model simulates a new dam volume as;

$$Vol_t = Vol_{t-1} + (rain-pevap)*area/1000000 \quad \ldots\ldots 4.44$$

where pevap is the monthly potential evaporation demand and t and t-1 relate to current and previous months respectively. When there is no rainfall for the particular month the dam's capacity is influenced by the evaporation only.

Water demand on the dam for irrigation purposes also influences the dam volume. If the monthly irrigation demand depth (Irrm) is positive the total irrigation demand (IrrDmd) volume is given as;

$$IrrDmd = Irrm * IrrAreaDmd \quad \ldots\ldots 4.45$$

and a new volume is computed in the model as;

$$Vol_t = Vol_{t-1} - IrrAreaDmd * Irrm \quad \ldots\ldots 4.46$$

4.12.2.2 Calibration principles

The reservoir parameters are not currently calibrated and there seems no justification to change this. The data needed are dam capacity, capacity-area curves and area above the dam. All these data are standard information for any large dam construction project and are quite easy to obtain. However most of the reservoirs that are found in most river basins are small farm dams and do not have this kind of information available. Methods are therefore required to get estimates of the data required. One way of doing this is through the use of remote sensing and GIS (Sawunyama, et al., 2006).

4.12.3 The main reservoir model parameters

The descriptions of the last two sections pertain essentially to the small farm dams which are found mainly on the tributaries of the river in a sub-basin. There is a different routine where a large dam exists on a river. It is from the small dam routines in that the inflows to the reservoir include all flow generated within the current sub-basin and from all upstream sub-basins. However there are similarities in the way the water balances calculations are performed for both the small dams and the large dams. A brief description of the parameters of the reservoir water balance model follows.

For all the sub-basins that have a reservoir in a system, the reservoir model makes use of an array model attribute that describes the compulsory input reservoir parameters at their outlets. The model attribute array made up of 14 rows for 14 different parameters and is used only where simulation for a reservoir is necessary. The parameters that are relevant for this simulation are;

i. Reservoir capacity which is given in million cubic meters (MCM).
ii. Dead storage of the reservoir which is given as a percentage of the reservoir capacity.
iii. Initial storage; this is the reservoir capacity at the beginning of the simulation period and is given as a percentage of the reservoir's capacity.
iv. A and B; these have the same interpretation as the ones described in section 4.12.2.
v. Reserve level; this describes a predetermined operation rule for the reservoir. It is given as a percentage of the reservoir's capacity. Up to 5 reserve levels can be defined for a reservoir. It determines a level of abstraction depending on the current capacity of the dam.
vi. Annual abstraction; this relates to the annual abstraction level for a given environmental reserve level and is given in MCM. This water is not available for abstraction downstream of the river. It is the flow required by the environment to maintain it at a predetermined level.
vii. This annual value will also be given with the associated monthly distributions at the 5 reserve levels given above. At times the reserve abstraction may be specified as a time series and in that case the parameters AR and BR (see (h) below) will be specified.
viii. Annual compensation flow; this is the annual downstream compensation flow released into the river from the reservoir. It is given in MCM.
ix. AR and BR; are the coefficients of the relationship between reserve and the volume. The relationship, which is non linear, is used to determine the amount of water released as environmental reserve depending on the storage level of the reservoir for any given month as follows;

$$\text{Reserve} = AR * \text{Volume}^{BR} \quad \text{..} \quad 4.47$$

where reserve is given as a percentage and volume as a percentage of the full capacity of the reservoir for a given month.

5 PARAMETER ESTIMATION METHODS FOR THE PITMAN MODEL

5.1 Introduction

This chapter outlines the physically-based parameter estimation procedures being proposed for the Pitman model. Estimations for two sets of parameters are described. These are, firstly, the soil moisture accounting, subsurface runoff and recharge parameters and, secondly, the soil surface infiltration parameters. The motivation for the development of the methods is multi-faceted. Firstly, the model is soundly based in conceptual hydrology and (taking into account spatial and temporal scale issues) the parameters should have physical meaning. This study therefore attempts to identify the conceptual linkages between the model formulation (and therefore its parameters) and physical basin properties as well as the way in which these affect hydrological processes. These conceptual linkages are then used to develop quantifiable relationships between measurable physical basin properties and the model parameters.

Secondly, if a practical parameter estimation approach is to be developed the information on physical basin properties must be available. There exist data collected by various agencies in the region which could possibly be used for hydrological modelling purposes though the data are not collected for this purpose. Such data relating to soil properties (FAO, 2003; Rawls et al., 1982), geological (Conrad, 2005) and topographical maps have a wealth of information that could be exploited for parameter estimation purposes. Meteorological data, such as rainfall (amount and distribution) and hydrogeological data (groundwater recharge, transmissivity and storativity) from various studies (e.g. DWAF, 2005; Xu and Beekman, 2003; Bredenkamp et al., 1995) have also been used for the current parameter estimation studies. It is, however, prudent to note at this juncture that these data are not available to the same level of detail or accuracy throughout the region.

Thirdly, while the availability of physical basin property data may present an existing limitation, new methods of collecting or processing such information (via satellite or through GIS applications) could improve the quality in the near future. Finally, a revised parameter estimation method could provide further direction and incentive for the collection and processing of physical basin property data.

For example, a Water Research Commission project has recently started where soils data relevant for use in, and improvement of, hydrological models are being collected with input from both soil scientists and hydrologists.

Modelling results using the Pitman model in selected basins in the southern African region have been very encouraging, but there is a need to further explore the regional application of this model. Hughes (1997) contends that there is potential for the regionalization of the model in southern Africa, but that further research is required to design more robust, less uncertain parameter estimation methods to realize this goal. While a number of techniques have been used for regionalization of the model, these have followed either of two formats as depicted in Fig. 5.1. Firstly, due to unknown parameter interdependences within the model as a result of the large number of parameters, there exists a number of equally good parameter sets (parameter sets 1 to n in Fig. 5.1A); the problem of equifinality discussed by Beven (1993). A set of basin properties is then collected to which the parameters are compared for regionalization purposes.

Some regionalization procedures have been qualitative in nature, in which parameter mapping based on some measure of basin similarity is used. In this case no relationships beyond descriptive analysis of the basin physical properties are used as the basis for regionalizing parameters and model simulations in ungauged basins (e.g. Midgley et al., 1994; Hughes et al., 2006). Other regionalization techniques have essentially followed a quantitative approach in which statistical relationships between optimized parameters (using gauged basins) and basin properties are developed. These relationships are then transferred to ungauged basins for estimation of parameter values (e.g. Hughes, 1982; Mazvimavi, 2003). The reliance of the regionalization process on the calibrated parameters introduces a measure of uncertainty in the model. The Pitman model is usually calibrated manually and therefore subjectivity may be a problem. For instance, several model users working on the same basin can quite easily produce different parameter sets giving equally good simulations if they concentrate on different components of the model. There are thus many difficulties in determining relationships between the parameters and basin properties for all the parameters sets (Fig 5.1A). This may make regionalization very difficult as it is difficult to choose the "best" parameter set for regionalization from the many possible sets.

The preceding chapter outlined the conceptual framework and physical interpretation of parameters that will be used in the current chapter to develop physically-based parameter estimation procedures. It is proposed to approach parameter value quantification and, subsequently, regionalization in a different way to the current procedures outlined above. If the parameter quantification process can be constrained using physical basin properties earlier on in the calibration process, the regionalization could be less uncertain. This implies that the use of physical basin properties in determining the parameter sets reduces the subjectivity in calibration and, therefore, equifinality, making it possible to obtain a basin specific, physically-based optimum parameter set. This is what is being proposed for the Pitman model through this study – a revised calibration process that takes into account the physical characteristics of the basin. With such a calibration procedure it is hoped that a single acceptable parameter set can be found for each basin regardless of the number or experience of model users.

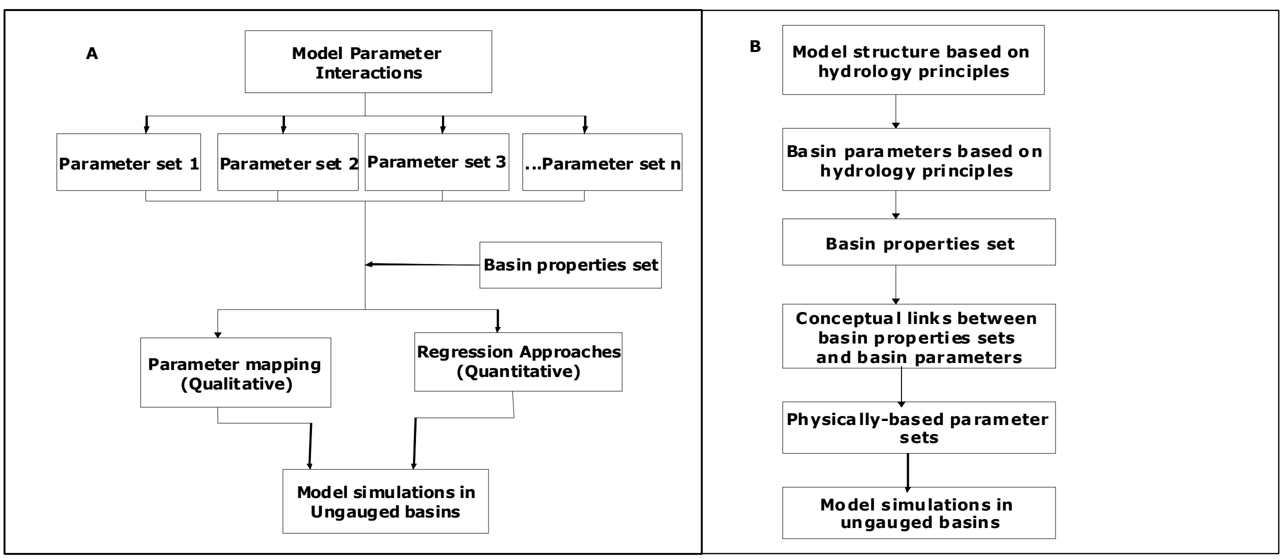


Figure 5.1 Approaches to parameter estimation and model regionalization used in the southern African region (A) and the proposed new approach (B).

The assumption of this approach is that both the model structure and basin parameters are based on sound physical hydrology principles. It is proposed that if a set of relevant basin physical properties data are available then it is feasible, using conceptual links between these properties and model parameters, to develop physically-based parameter estimation procedures following the path depicted in Figure 5.1B. The focus in this study is on the main runoff generation parameters, while it is assumed that the approach could ultimately be extended to the full parameter set of the model.

Within the following two sections reference is made to a simple Delphi computer program which has been established to provide default estimates of the relevant physical basin properties and the parameter values. The information requirements of this approach have been kept to a minimum to facilitate its use. There could be, however, better methods of estimating some of the basin properties if more information is available. These are referred to when appropriate and the method used in any specific basin should be that which is likely to generate the most representative value of any basin property.

5.2 Soil moisture accounting, subsurface runoff and recharge parameters

A case was made for the physical interpretation and estimation of the soil moisture parameters in the previous chapter (sections 4.6 and 4.7). The physical basin property data required for the purpose of the estimation are dictated by the conceptual framework used for interpretation of these parameters in the structure of the model. This section will address the maximum soil storage (ST), subsurface runoff (FT and POW) and groundwater recharge (GW, GPOW) parameters. As explained in the previous chapter the maximum moisture storage (ST) parameter of the Pitman model is assumed to be made up of two components; the soil component and the unsaturated zone component (i.e. $ST = ST_{soil} + ST_{unsat}$). Subsurface runoff is also assumed to be generated separately from these two components (i.e. $FT = FT_{soil} + FT_{unsat}$).

5.2.1 Estimating ST_{soil}

ST_{soil} represents the soil storage depth (mm) at saturation and is very important in hydrology as it represents the immediate store of infiltrated rainfall before it is lost to either evapotranspiration or to percolation and runoff. The maximum amount of moisture of the 'soil' component (ST_{soil}) is estimated by the following equation;

$$ST_{soil}\ (mm) = POR\ (\%) * VVAR\ (\%) * Soil\ Depth\ (m) / 10 \qquad 5.1$$

POR represents the soil porosity and therefore a measure of the moisture holding capacity and VVAR represents a correction factor for vertical variations in porosity. Figure 5.2 illustrates that the default estimate of porosity used in this study is primarily based on the soil texture class. Many previous studies have related porosity to the percentage distribution of sand, clay and silt within different texture classes (USDA, 1969; Rawls et al., 1982; Schulze et al., 1985). In this study, 5 texture classes have been used for the default approach and the assumed porosity (%) values are given in column 6 of Table 5.1. Figure 5.2 shows that the final mean porosity for the basin can be made up of area weighted averages of the 5 texture classes. Soil depth is estimated from the percentage areas of the basin occupied by three main topographic units (upper slope, mid slope and valley bottom – see Fig. 5.2) and the average soil depths associated with them. The final soil depth used in equation 5.1 is the area weighted average.

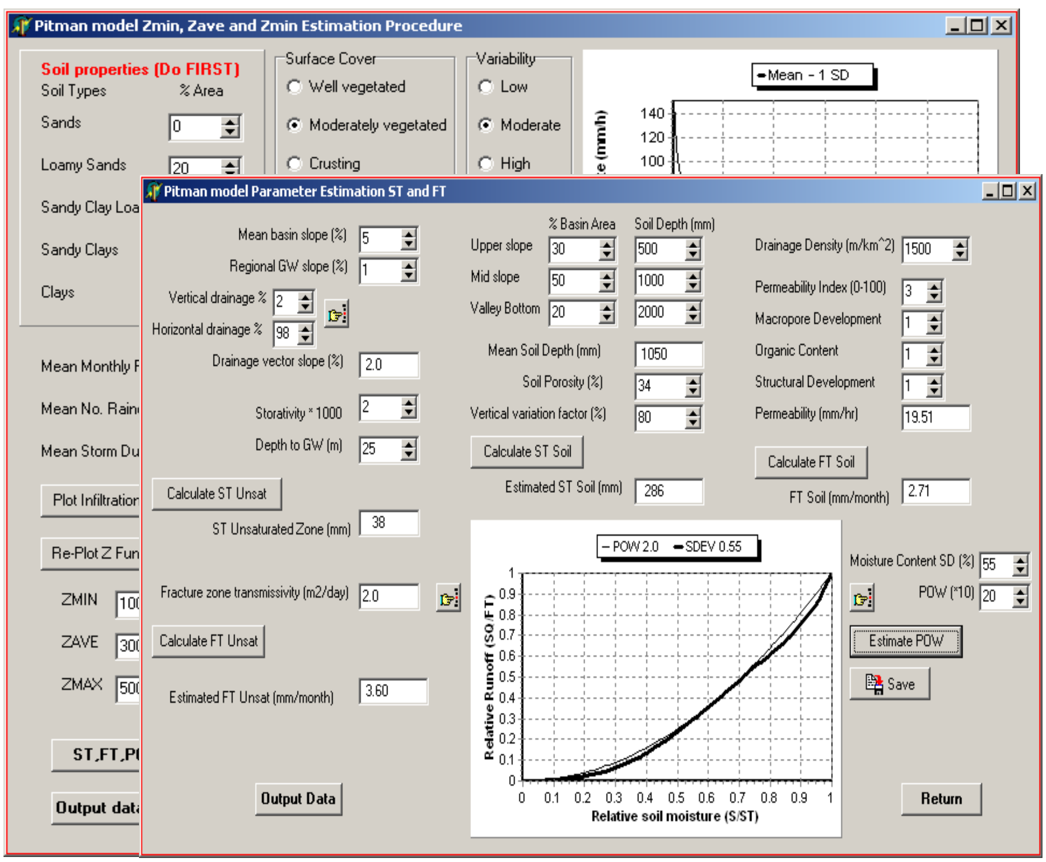


Figure 5.2 Illustration of the default basin property and parameter estimation program.

Table 5.1 Soil texture classes according to USDA (1969), based on percentage volumes of sand, silt, clay and quartz content.

Texture Class	Sand (%)	Silt (%)	Clay (%)	Quartz (%)	Assumed porosity (%)
Sand	92	5	3	92	42
Loamy sand	82	12	6	82	40
Sandy loam	58	32	10	60	Not used
Loam	43	39	18	40	Not used
Silt loam	17	70	13	25	Not used
Silt	10	85	5	10	Not used
Sandy clay loam	58	15	27	60	33
Clay loam	32	34	34	35	Not used
Silty clay loam	10	56	34	10	Not used
Sandy clay	52	6	42	52	32
Silty clay	6	47	47	10	Not used
Clay	22	20	58	25	39

The main sources of uncertainty are associated with the spatial variation of both depth and texture and the most appropriate method that is used to specify basin averages. This issue could be resolved by subdividing the basin into sub-units that have relatively uniform soil depths and texture. These are expected to be closely related to topography in the same way that the South African land type maps are developed (SIRI, 1987). The final ST_{soil} value would then be quantified as an area weighted average of the ST_{soil} values in all the units of the basin. This type of analysis is easily achieved using GIS software.

5.2.2 Estimating ST_{unsat}

The unsaturated zone between the water table and the soil zone is quite difficult to characterize, given that there are gaps in our understanding of the water transfer processes that operate there. The assumption made in this study is that water percolating downwards in the unsaturated zone will have two directional components; a vertical one contributing directly to recharge of the saturated ground water zone and a lateral one that could contribute to the re-emergence of subsurface water at a spring or seep. The important issue is that these springs or seeps occur at elevations above the regional ground water level. The lateral component could be caused by flow in horizontal fractures or through perched aquifers associated with layers of lower permeability. The vector result of these two components is referred to here as the drainage vector slope (VS in Fig. 5.3), which is estimated in the default procedure using % values for the vertical and horizontal components. Approximate estimates for different geological conditions also form part of the estimation procedure (Fig. 5.4). A high vertical component results in a steep drainage vector which would be prevalent in permeable rocks without impermeable layers or lenses that may induce lateral flow. Low VS values will be found in situations where there are many horizontal fractures compared to vertical fractures, or where impeding layers exist in otherwise permeable material. For unsaturated flow to re-emerge as spring flow, VS must be less than the mean basin slope (Fig. 5.3).

The ratio of the volume that lies between the basin surface slope (BS in Fig. 5.3) and the drainage vector slope (VS) to the total unsaturated volume represents the proportion of the unsaturated zone that can contribute to unsaturated flow. The area between the drainage vector slope and the ground water slope (GS) will not be able to contribute to unsaturated flow at the surface, but will contribute to aquifer recharge. Simple geometry from Fig. 5.3 suggests that the ratio of these two areas can be calculated from:

Ratio = [Tan(BS) – Tan(VS)] / [Tan(BS) – Tan(GS)] 5.2

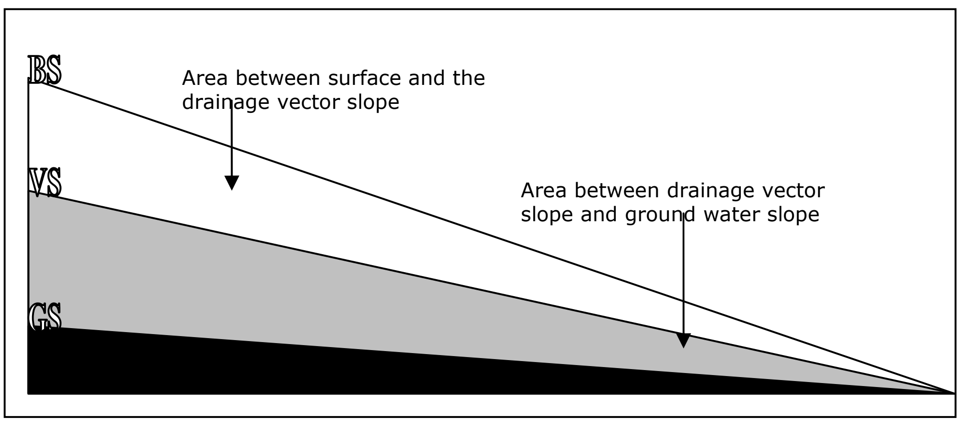


Figure 5.3 Conceptualization of the subsurface drainage that determines the interflow process from the unsaturated zone.

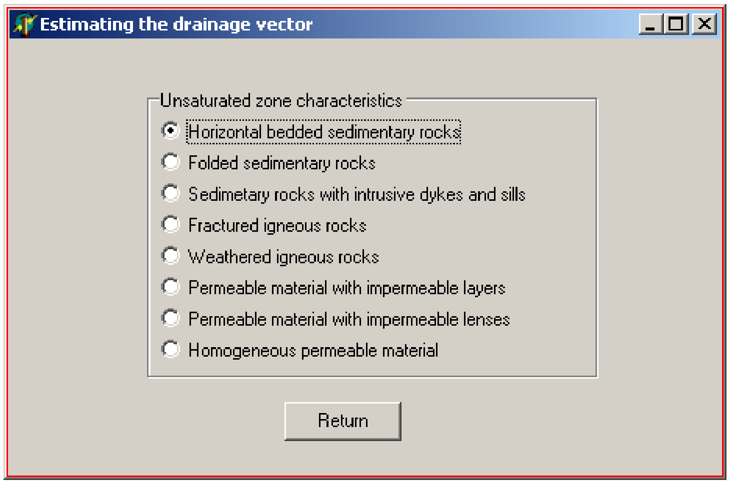


Figure 5.4 Default estimation approach for the drainage vector slope.

If the total unsaturated zone potential storage (mm depth) is expressed as the product of the mean depth to ground water (DGW m) and the storativity (S) of the unsaturated zone material, then the final estimate of ST_{unsat} becomes:

If BS > VS then

ST_{unsat} (mm) = DGW * 1000 * S * Ratio 5.3

If BS ≤ VS then

ST_{unsat} (mm) = 0 .... 5.4

With respect to the use of this estimation approach, several important issues have to be considered. The mean depth to ground water may be available from regional borehole surveys, but it is important to recognize that any available values must be consistent with the conceptualization of the estimation approach and are not biased by preferential borehole locations. The storativity value used must represent the component of the unsaturated zone that can contribute to 'unsaturated' flow. In a fractured rock situation this would therefore represent the fracture storativity only (excluding the rock matrix storage). If the drainage vector slope is close to the ground water slope, almost all of the unsaturated zone can contribute.

The information on depth to ground water needed for this estimation may be reasonably accurate in areas where comprehensive borehole drilling records exist. Estimations for areas without this kind of information may introduce some uncertainties. Obtaining representative values of storativity may also be a problem in some areas and may introduce a further source of uncertainty.

5.2.3 Estimating FT_{soil}

Subsurface lateral and vertical drainage are known to occur at different moisture contents. Field capacity defines the volume of moisture that a particular soil is capable of holding against the force of gravity. In a purely Darcian flow context no significant water movements occur below field capacity, while at higher moisture contents vertical drainage can occur. Significant volumes of lateral flow only occur close to saturation levels. However, spatial variations in soil characteristics at all scales and the well-documented presence of macro-pores (e.g. Inoue, 1993; Greco, 2002) suggest that simplistic applications of Darcian flow concepts are frequently inapplicable. The implication is that sub-surface lateral flow can occur within a basin over a wide range of average basin moisture contents. This is implicit in the Pitman model 'soil' moisture runoff generation component.

FT_{soil} is the maximum subsurface outflow when the basin's soils are at saturation and is assumed to occur through the banks of the active channel. At saturation, therefore, the whole stream channel is active and the average soil depth gives an estimate of the depth of the channel through which water flows into the river. The total contributing channel length is estimated from the basin drainage density. Thus, shrinkage of the drainage density should reduce the volume of subsurface flow. Since both banks are active, the estimation equation is multiplied by 2. The contributing area (CA, in $km^2\ km^{-2}$) is given by:

$$CA = 2 * DD\ (km\ km^{-2}) * \text{soil depth (m)} / 1000 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ 5.5$$

where DD is the basin's drainage density. The soil depth value used should be based on the soil depths in the lower topographic units of the basin. The drainage density is a measure of channel length and can be estimated from topographic maps. The calculation of drainage densities included all potential drainage lines (identified by contour convergence) that are assumed to receive flow under conditions of basin saturation. While this makes the drainage densities higher than the use of 'blue' lines, it was assumed to be reasonably realistic under saturation circumstances when many seasonal streams form. Comparisons of drainage density measured from 1:250 000 and 1:50 000 maps showed that estimates from the more detailed maps were about three times the ones from the 1:250 000 maps. This scaling factor has been used within this study where a more rapid estimate has been based on 1:250 000 scale maps.

The monthly depth of interflow from the soil (FT_{soil}, in mm month^{-1}) was thus assumed to be adequately explained as a function of CA, saturated hydraulic conductivity of the basin soils, K (m d^{-1}) and the mean basin slope (BS) and expressed as follows;

FT_{soil} = CA * K * BS * 30 * 1000 ……………………………………………………………….. 5.6

The estimation approach for K is illustrated in Fig. 5.2 and is based on area weighted soil texture classes plus some adjustments to account for macro-pore development, organic content, structural development and sand grade. Cosby et al. (1984) suggested typical means and ranges of hydraulic conductivity values for different soil types and these were used as a guide in this study. The actual values of K used in this study are based on the various factors that operate on a basin scale using the following relationship:

K (m/day) = $e^{(PI*0.55 - 0.054)}$ …………………………………………………………………….. 5.7

Where PI is a permeability index value estimated from soil characteristics and is given by:

PI = M + 0.5 * (F+G+H) + K ………………………………………………………………………. 5.8

where

M = 0.09A + 0.05B +0.02C + 0.015D + 0.01E ……………………………………….. 5.9

and A to E are percentage areas of the basin covered by sandy (A), loamy sand (B), sandy clay loam (C), sandy clay (D) and clay (E) soils, while F, G and H are assumed to vary from low (0) to high (2) and represent the level of macro-pore development (F), the organic content (G) and the structural development of the soil (H). K represents the sand grade of the soil, which has been fixed at an index value of 1 in this study. This estimation procedure has been adopted from the methods used for the VTI model (Hughes and Sami, 1994). The resulting values of hydraulic conductivity (Table 5.2) correspond quite well with those given in Cosby et al. (1984) and Rawls et al. (1982). Using the default estimations of K values of FT_{soil} were calculated and are given in Table 5.3.

Table 5.2 Comparison of values of hydraulic conductitivity (in m d^{-1}) by three estimation methods (F, G, H = 1 in column 4).

Texture class	Cosby et al. (m d^{-1})	Rawls et al. (m d^{-1})	Using Permeability Index (m d^{-1})
Sand	4.03	5.04	4.98
Loamy sand	1.22	1.47	1.23
Sandy clay loam	0.38	0.10	0.39
Sandy clay	0.62	0.03	0.43
Clay	0.08	0.01	0.08

Table 5.3 Results of default estimation procedure for FT_{soil} (mm $month^{-1}$).

Drainage density	Low = 1.5			High = 2.5		
Soil Depth	Shallow	Moderate	Deep	Shallow	Moderate	Deep
Texture and hill slope gradient						
Loamy Sand/5%	2.0	5.0	10.8	3.3	9.0	18.1
Loamy Sand/10%	4.0	10.8	21.7	6.6	18.1	36.2
Loamy Sand/20%	8.0	21.7	43.4	13.3	36.2	72.3
Sand Clay Loam/5%	0.4	1.0	2.1	0.6	1.7	3.5
Sand Clay Loam/10%	0.8	2.1	4.2	1.3	3.5	6.9
Sand Clay Loam/20%	1.5	4.2	8.3	2.5	7.0	13.9
Clay/5%	0.2	0.6	1.2	0.4	1.0	2.0
Clay/10%	0.4	1.2	2.4	0.7	2.0	4.0
Clay/20%	0.9	2.4	4.8	1.5	4.0	8.0

Note: The following soil depths are assumed: Shallow= 0.25m, moderate= 0.75m, deep= 1.5m. Indices for macro-pore development (F), organic content (G) and structural development (H) are all equal to 1.

5.2.4 Estimating FT_{unsat}

Estimating the outflow from the unsaturated zone (FT_{unsat}) is by far a greater challenge. This is mainly because the physical concepts of subsurface runoff generation from this zone are not very well defined. For instance, there is no general consensus on the processes which occur in the unsaturated zone (below the root zone and above the ground water table). There is also limited documentation of the typical hydraulic conductivities of fracture zones. Figure 5.3 represents a conceptual diagram that is independent of the actual processes occurring. The lateral component contributing to the drainage vector may be the result of water flowing in horizontal, or near horizontal, fractures. It may also be a result of a series of overlapping layers of material with low permeability creating perched water tables and allowing lateral saturated flow to develop. The estimation approach adopted assumes either saturated flow in the fracture zones or a perched water table and is based on defining a representative transmissivity (T in m^2 d^{-1}):

$$FT_{unsat}(mm) = 2 * DD * T * VS * 30 / 100 \quad \text{.......} \quad 5.10$$

The quantification of a representative value for the transmissivity is the major source of uncertainty. While transmissivity values in fractures can be very high (Razack and Lasm, 2006), the fractures represent a variable but generally small proportion of the total volume of the unsaturated zone. This will depend on the degree of fracturing and the connectivity of individual fractures. In a perched aquifer situation, estimation of FT_{unsat} will depend on the transmissivity of the more permeable layers as well as on the number and geometric arrangement of the impermeable layers. It should be further emphasized that the transmissivity value used must represent the sub-basin as a whole, which accounts for variability in the geology. The values currently used within the default estimation program vary from 0.5 to 5 m^2 d^{-1}, however, further confirmation of appropriate values is required.

The drainage density used in the default estimation equation is the same as for FT_{soil}. However, it is also possible that the length of the channel that receives flow from the unsaturated zone could be less than the length receiving flow from near surface soil saturated flow. It is difficult to offer generic guidelines as individual basins may experience different conditions.

5.2.5 Estimating POW

The power (POW) of the relationship between subsurface outflow and the volume of moisture in a basin is assumed to be made up of the two components associated with the soil water and the unsaturated zone runoff. POW represents the shape of the relationship that determines reduced runoff (relative to the maximum) as the moisture contents of the soil and unsaturated zones decrease. This reduced runoff may be caused by reduced areas of saturation and therefore reduced contributing area, or it may be caused by reduced rates of runoff. In the soil zone the relationship is likely to be mainly influenced by patterns of moisture redistribution following rainfall events and how these patterns affect the distribution of saturated zones. The redistribution will be influenced by such processes as evapotranspiration, and vertical and lateral drainage.

Geology, topography, vegetation cover, soil type and texture will all influence patterns of moisture redistribution within a basin. It is therefore reasonable to suggest that, for any given mean basin moisture content (S), the spatial variation could be represented by a frequency distribution. At the extreme ends of the moisture content spectrum, i.e. when the basin is either very dry or close to saturation, this variability must be low. The variability would be highest at moderate moisture contents. Given detailed field observations the spatial variation of moisture content could be adequately defined for a range of basin mean moisture contents. However, in the absence of detailed field data, a simpler approach was adopted based on the probability distributed principle of Moore (1985) and similar to the procedures used within the VTI model (Hughes and Sami, 1994).

The concept is illustrated in Fig. 5.5. The four lines represent cumulative Normal distribution frequency curves for mean basin moisture contents of 0.2 to 0.8, each having a different standard deviation. If a relative moisture content of 0.9 is assumed to represent the threshold for lateral flow, Fig 5.5 indicates that the percentage of the basin area contributing to runoff would vary from 0% (at mean of 0.2) to 60% at a mean of 0.8 (triangle symbols). If a method of estimating the variation in the standard deviation with mean moisture content can be found, it follows that a relationship between mean moisture content and relative runoff (i.e. runoff relative to the maximum at full basin saturation) can be developed. The principles of such a method should be that the standard deviation will be at a maximum at moderate moisture contents and at a minimum for both low and high moisture contents. The approach adopted uses quite arbitrary equations to achieve these principles and is based on a SDEV parameter that is assumed to vary with basin properties:

If RAT > 0.75

$$SD = (1.1 - RAT) * SDEV / (1.1 - 0.75) \qquad 5.11$$

If RAT ≤ 0.75

$$SD = (RAT + (0.75 - RAT) * 0.2) * SDEV / 0.75 \qquad 5.12$$

where RAT = mean relative moisture content,

SDEV = maximum standard deviation, and

SD = standard deviation at RAT.

The relative runoff is then calculated from the proportion of the frequency distribution that exceeds a relative moisture content of 0.9. A correction factor is sometimes necessary to ensure that the relative runoff is 1.0 at a mean relative moisture content of 1.0. The correction factor used simply scales all the relative runoff estimates (for the full range of mean moisture contents) proportionately.

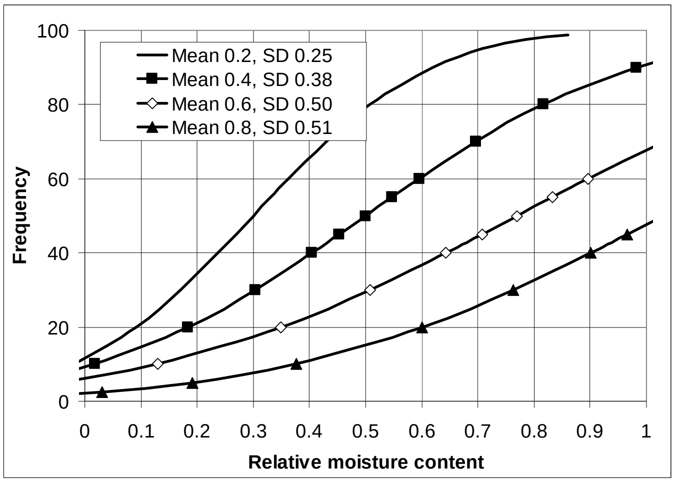


Figure 5.5 Illustration of the concept of using a frequency distribution to describe the spatial distribution of soil moisture for different mean moisture contents.

The resulting relationship between mean relative basin moisture content and relative runoff is then identical to the format of the Pitman model 'soil' moisture runoff function if this is expressed in non-dimensional terms (i.e. S/ST_{soil} for the horizontal axis and Q/FT_{soil} for the vertical axis – see Fig. 5.6). It is assumed that the standard deviation (SDEV) at a mean moisture content of 0.75 can be established from the physical attributes of the basin. Low values of SDEV are expected when there is little spatial variation in moisture content, which may occur in areas of low topography and poorly drained soils. Most variations will then be a result of variations in evaporative loss. High values of SDEV are expected in steep topography with well drained soils on the hill slopes and less well drained soils in the valley bottoms. After rainfall events the well drained soils will dry out, contributing to additional moisture content in the lower slopes and maintain soil wetness in those areas. These concepts are represented in Fig. 5.6

and appropriate values of POW in the Pitman model function ($Q/FT = (S/ST)^{POW}$) have been manually calibrated to reproduce similar shaped curves.

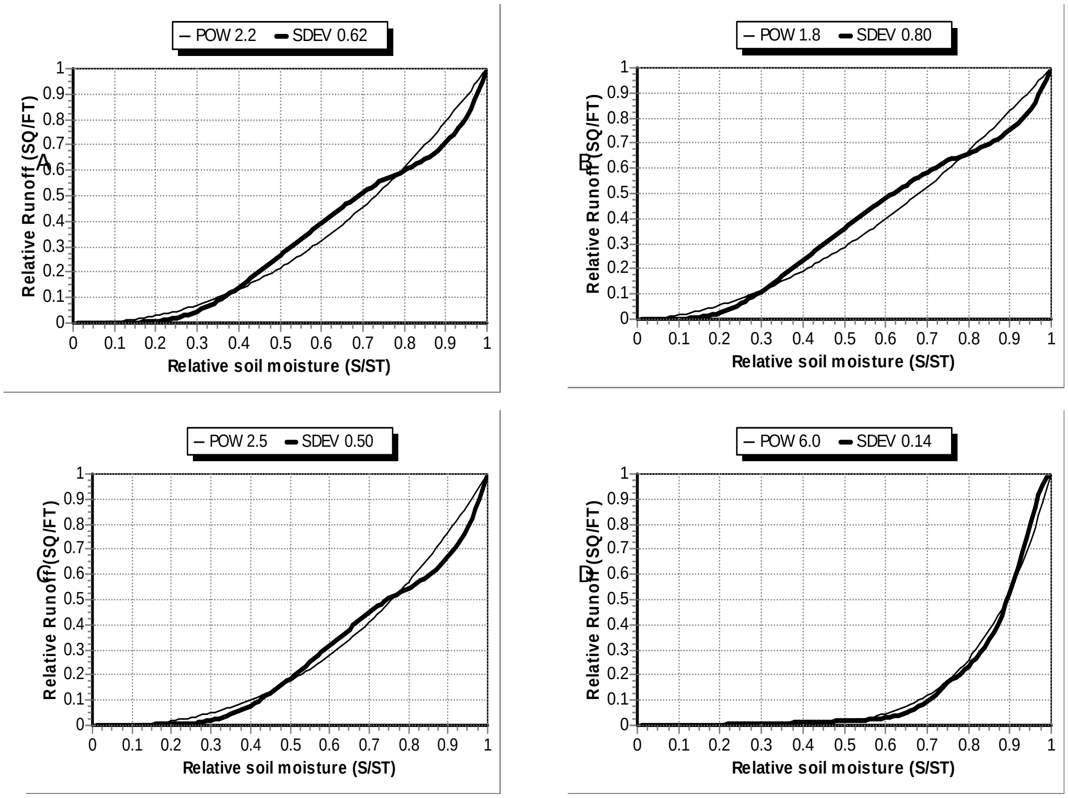


Figure 5.6 Runoff-moisture content relationships for four conditions (defined by the moisture distribution parameter, SDEV). The basin conditions represented are steep slopes and well drained soils (A), moderate slopes and moderately well drained soils (B), moderate slopes and moderately poorly drained soils (C) and gentle slopes and poorly drained soils (D).

The previous paragraphs have ignored the contribution of the unsaturated zone and this component of the relationship is more difficult to assess. It is assumed to be related to a decrease in the number of saturated fractures as well as a reduction in the drainage density of channels receiving spring flow as the moisture content (S_{unsat}) reduces. In the absence of a better defined approach the shape of the unsaturated curve is given by a simple power function:

$$Q_{unsat}/FT_{unsat} = (S_{unsat}/ST_{unsat})^2 \quad \text{5.12}$$

where Q_{unsat} represents the unsaturated zone runoff at unsaturated zone mean moisture content of S_{unsat}.

The full estimation approach is to generate the two curves separately (soil and unsaturated zones) and then adjust both to ensure that the ordinates range from 0 to 1. The adjustment is based on the relative contributions to total runoff of the two zones (i.e. FT_{soil} and FT_{unsat}). Figure 5.7 illustrates the effect of excluding and including FT_{unsat} in the estimation. There is not a large difference in this case but it is expected that in areas where the contribution of the unsaturated zone is high its exclusion would lead to errors in the estimation of POW.

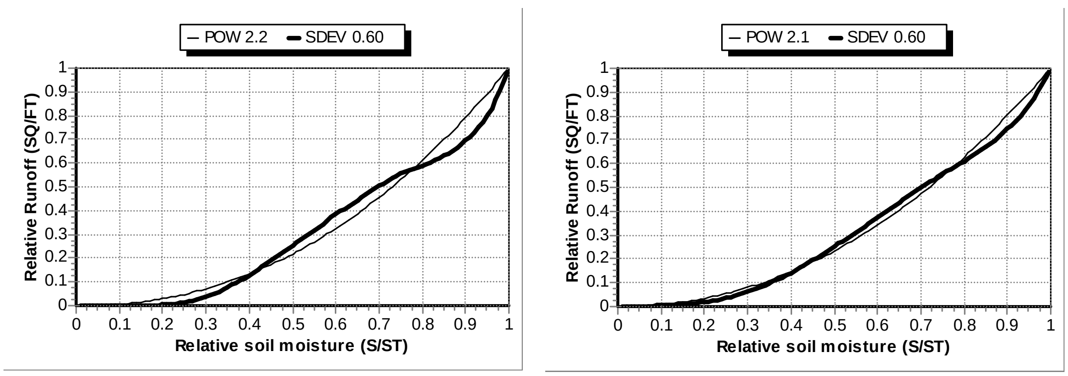


Figure 5.7 Runoff-moisture content relationships for the same basin without FT_{unsat} (left side) and with FT_{unsat} (right side). The value of FT_{unsat} is 6.1mm and FT_{soil} is 8.0mm.

A single value of S (mean moisture content) is calculated in each time step of the model and represents both the soil and unsaturated zones. In practice the mean moisture contents of these zones would not vary with the same pattern (the unsaturated zone would tend to lag behind changes in the soil zone). However, to incorporate such a modification would require substantial changes to the model structure, which is not the purpose of this study.

5.2.6 Estimating GW and GPOW

Estimating the value of GW is difficult as it involves the complexities of vertical drainage through the total unsaturated zone. The approach to estimating GW and GPOW could follow similar principles to those used for FT_{soil} and POW. There are, however, existing estimates of mean annual recharge available for some southern African basins which can be used to guide the calibration of GW (see section

4.7.3). GPOW will be similarly calibrated and results compared against observed low flows where available. These approaches are considered to be adequate at this stage of the development of the parameter estimation procedures.

5.3 Soil surface infiltration parameters

As pointed out in the previous chapter these parameters control the absorption rate at the surface, the volume of water entering the moisture store reservoir and the depth of infiltration excess flow generated within a particular basin. Ponding occurs when the rainfall rate is greater than the infiltration capacity of the soil and is an important aid to the process of infiltration at the basin scale. However infiltration rates tend to decrease with time under ponded conditions and will approach the saturated hydraulic conductivity (K_{sat}) of the soil due to the weakening of the energy gradient in Darcy's law as the soil gets wetter. Under non-ponded conditions infiltration rates will vary with the rates of the rainfall input. An array of factors influences the process of infiltration, chief among them being soil properties (both physical and hydraulic) and antecedent moisture conditions and these factors are used here in developing the new physically-based estimation procedures for the soil surface infiltration parameters.

The approach taken for the design of a physically-based procedure makes use of both basin surface and hydro-meteorological factors. The basic tenet of this approach is to use soil properties to define the parameters of a modified form of the Kostiakov equation (Hughes and Sami, 1994), basin hydro-meteorological characteristics to disaggregate monthly rainfall and to apply the infiltration equation to estimate surface runoff for a range of monthly rainfalls. The parameters ZMIN, ZAVE and ZMAX (section 4.4.2) of the surface runoff model algorithm are then manually fitted to match the infiltration equation based estimates of runoff for different monthly rainfalls.

5.3.1 Generating runoff using the infiltration excess function

This procedure is based on the use of a variation of the Kostiakov equation (Kostiakov, 1932) to estimate surface infiltration rate as follows:

$$\text{Infiltration rate (mm h}^{-1}\text{)} = k * C * T^{k-1} \qquad \text{.......... 5.12}$$

where k and C are parameters and T is cumulative time in minutes from the start of the storm. The mean values of the parameters and their assumed spatial variability (expressed as the standard deviation of a log-normal distribution) are estimated from soil texture properties and surface cover. The top part of Figure 5.8 illustrates the approach as well as a graphical representation of the infiltration equation and its variability. The details of this approach can be found in Hughes and Sami (1994). The approach incorporates the principle of spatial variability in infiltration rates over the sub-basin and allows for this variability in estimating the surface runoff at any specific rainfall rate.

In order to apply the infiltration function to monthly rainfall totals it is necessary to first disaggregate the monthly rainfall. Within the Pitman model, monthly rainfalls are disaggregated into four periods (see section 4.2.1) and the same equations are used here but disaggregating into 30 periods (i.e. approximately 1 day per period). This disaggregation process generates rainfall on every day of the month, which is not a very realistic distribution to use with the infiltration function. A parameter representing the mean number of rain days expected in the basin is used to aggregate some of the initial daily rainfall estimates and leave some days with zero rainfall. The assumption is that any daily rainfall must be greater than the square root of the ratio of monthly rainfall total divided by the mean number of rain days. This is loosely based on the probability of occurrence of a rain day (de Groen, 2002) which in this study is taken as a day with rain above a certain threshold (as opposed to a day with recorded rain). This threshold is defined by the total monthly rainfall and the mean number of rain days. Such an approach is deemed adequate for the disaggregation of monthly rain. Worldwide estimates of the number of rain days per month can be obtained from the IWMI Online Climate Summary Service (www.lk.iwmi.org/WAtlas/AtlasQuery.htm).

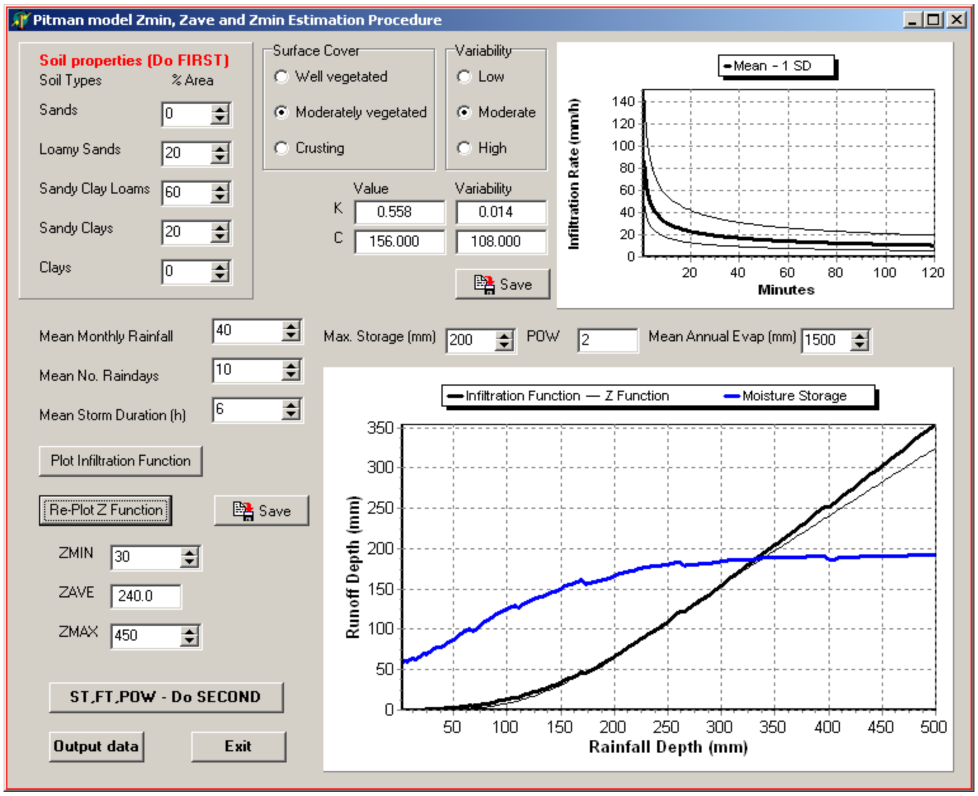


Figure 5.8 Illustration of the use of the default estimation procedures for the surface runoff parameters (ZMIN, ZAVE and ZMAX).

The daily rainfalls are further disaggregated into 5 time periods based on a parameter representing the expected mean storm duration, MSD (h) using:

RINT = rain * stvar (k) * 60 / MSD ….…………………………………………………. 5.13

where RINT is the rainfall intensity (mm h-1), rain is the daily rainfall (mm), and stvar is a distribution constant for each of the 5 time periods (0.045, 0.184, 0.383, 0.300, and 0.088 for time periods 1 to 5 respectively).

These rainfall rates are then compared with the frequency distribution of infiltration rates at the appropriate time since the start of the storm to generate an initial estimate of surface runoff in exactly the same way as applied within the VTI model (Hughes and Sami, 1994). During the early stages of attempting to apply this approach to estimate the Pitman model surface runoff parameters, it was noted that the effects of saturation excess surface runoff (runoff generated from rain falling on saturated parts of the sub-basin) could not be ignored. Consequently, a simplified water balance estimation has been included with the soil moisture storage being updated during the month (it is assumed to start at 30% of total saturation).

The Pitman model soil moisture runoff function (using previously determined values for ST_{soil} and FT_{soil}) is applied as part of this water balance estimation together with a rough estimate of evaporative losses. The value of the relative soil moisture runoff estimate (Q/FT_{soil}) is used to estimate the proportion of the basin that is saturated at the surface (from which all the rainfall will contribute to runoff) and which must be excluded from the part of the sub-basin that can potentially generate infiltration excess runoff. The runoff generated in all time steps within the month is accumulated to give the total monthly runoff and plotted against the monthly rainfall (see the lower part of Fig. 5.8).

The parameters (ZMIN, ZAVE and ZMAX) of the Pitman model triangular surface runoff (absorption) function are then adjusted and the function plotted as a cumulative curve to be similar to the infiltration function results (Fig. 5.8). It should be noted that it is generally not possible to achieve a fit between the two curves for the whole range of rainfall depths. However, experience of the method suggests that it is usually possible to get a reasonable fit over the range of rainfalls for a sub-basin and that the resulting values of the model parameters are consistent with values typically used in previous applications of the model.

While the estimation procedure used incorporates the concepts of both infiltration excess as well as saturation excess runoff, the Pitman model does not and in the model the surface runoff estimations are explicitly independent of moisture storage conditions. In developing the parameter estimation approach, the issue of saturation excess runoff could not be ignored and it is assumed that this is related to the difference in time scales used in the estimation procedure compared to the Pitman model algorithm. It is possible therefore that the Pitman model surface runoff algorithm is implicitly accounting for saturation excess runoff despite not being directly related to the simulated soil moisture level.

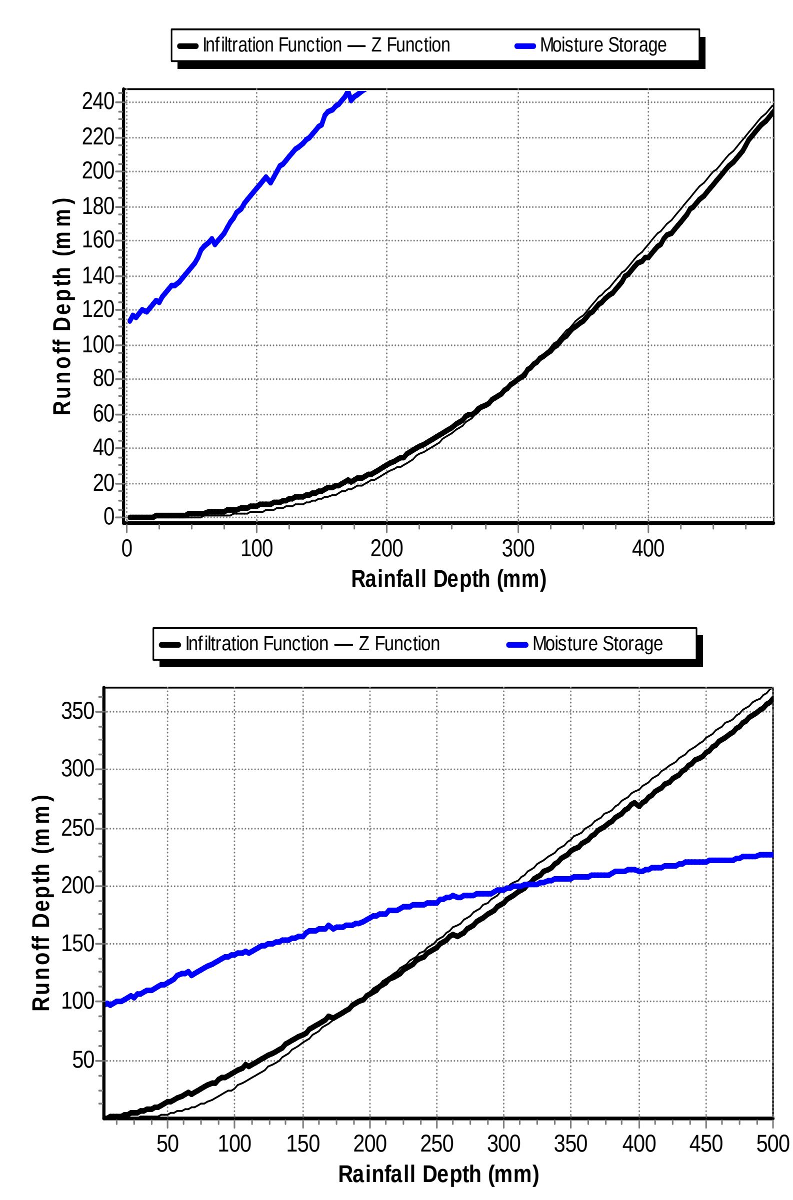


Figure 5.9 Illustration of the estimation of ZMIN, ZAVE and ZMAX for two situations. The diagram on the left represents sandy soils of moderate depth with ZMIN = 30, ZAVE = 415 and ZMAX = 800. The diagram on the right represents crusted clay soils of moderate depth with ZMIN = 0, ZAVE = 120 and ZMAX = 350.

6 RESULTS

6.1 Introduction

This section presents the major findings of the study. The results are based on a total of 71 basins chosen from the southern African region that were investigated. Appendix 1 and Appendix 2 show the physical basin property data for all the basins, while examples are used within this section for illustrative purposes. In general, the choice of the basins was mainly influenced by the availability and quality of rainfall and observed streamflow data. The length of the modelling periods used for the different basins was similarly influenced by the rainfall and runoff data. However, the basins were also chosen to reflect the diversity of typical physical basin properties (i.e. soil texture classes, soil hydraulic properties, geological and topographical conditions, climate and runoff regimes) obtaining in the region. This section provides a summary of the major physical characteristics of the basins, the parameters estimated by the revised methods and the results of simulations using these parameters.

6.2 Basin characteristics

6.2.1 Climate, relief, geology and soil

The climate of the southern African region is very diverse with arid conditions being experienced in the western parts, in countries as Botswana and Namibia, and more humid temperate sub-tropical conditions in the south-western and north-eastern parts of South Africa, northern and western Mozambique, eastern and central Zimbabwe, north-western Zambia and central Malawi. The mean annual precipitation (MAP) and the mean annual potential evapotranspiration (MAE) were used as rough indicators of climate (Fig. 6.1). The two sub-basins from Botswana that are part of the Metsemothlaba River system and the south eastern part of the Berg basin (e.g. the Sout River system, G50G) in South Africa were among the basins chosen to test the parameter estimation methodology in the arid parts of the region. From the more humid parts of the region the Kafue system in Zambia, the headwater basins of the Pungwe River (F14, F22), the Zonwe River (F10) and the Budzi River system (F18) in Zimbabwe and the Sabie (e.g. X31A, X12J) and the Tugela (e.g. V20A, V70D) in South Africa were chosen for the study. All the other basins fall within these two extreme climatic conditions.

A. Mean annual evapotranspiration (MAE)

B. Mean annual precipitation (MAP)

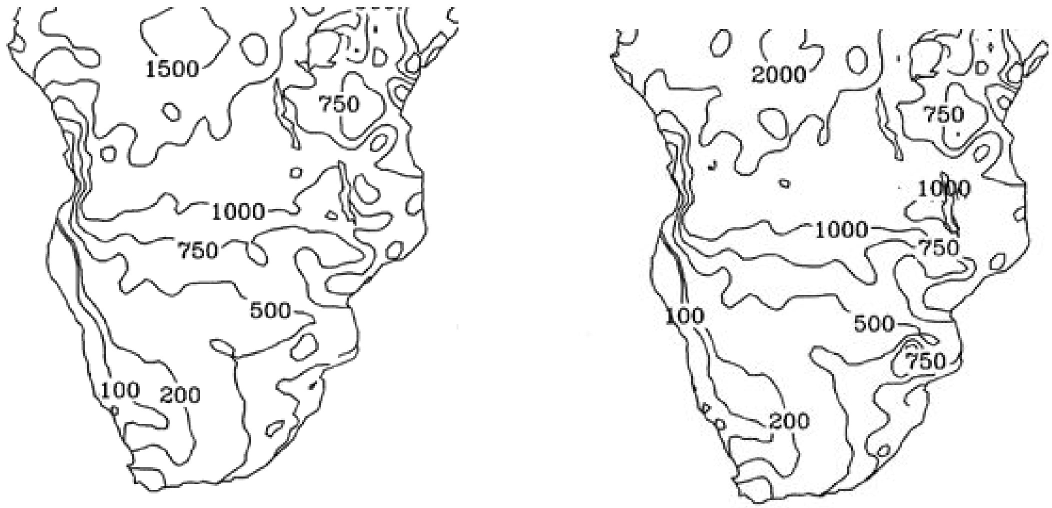
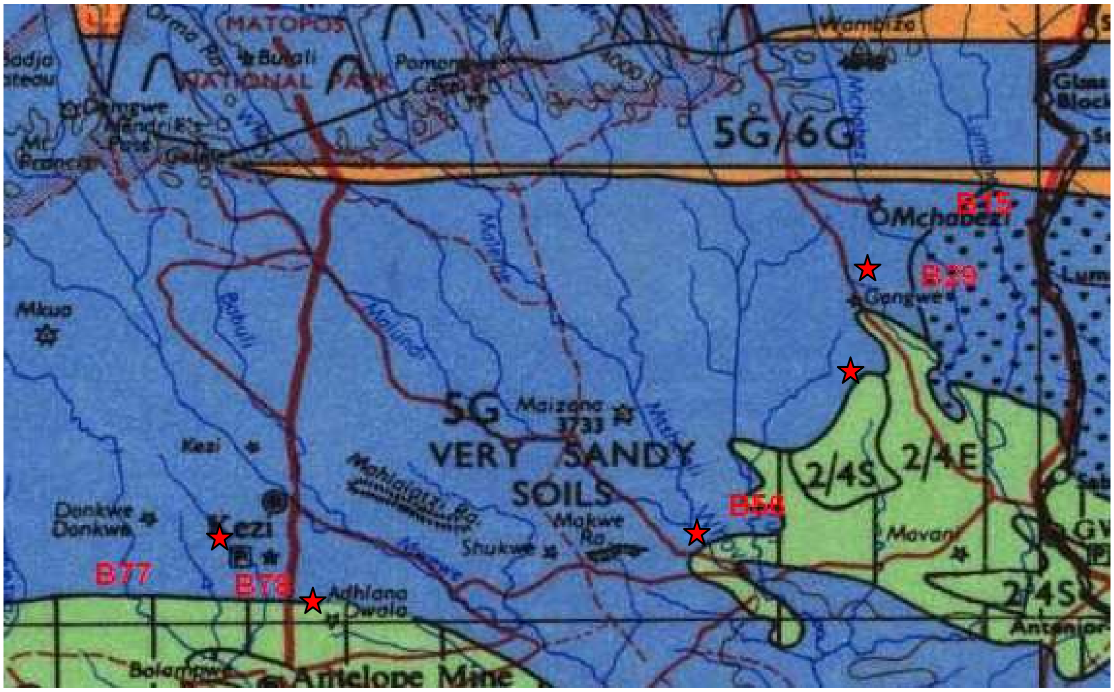


Figure 6.1 The distribution of mean annual potential evapotranspiration (MAE) and mean annual precipitation (MAP) over southern Africa. The MAP is a 40 year average for the period between 1950 and 1989, (Nicholson et al., 1997).

The relief is also equally varied from relatively flat, near sea level areas (e.g. E67 in Mozambique) through undulating topography (e.g. E63, E136 and F1 in Zimbabwe, 2421 and 2411 in Botswana) to steep topography basins (e.g. K20A, V70D, G10E and X31A in South Africa). The information on relief was obtained from maps (either 1: 50 000 or 1: 250 000) of the basins where these were available. The maps for the South African basins were available at both scales.

Geologically, most of the region is underlain by an assortment of Precambrian formations which are quite deeply weathered, or substantially fractured, rocks of volcanic and metamorphic origin and also large portions of sedimentary rock formations. From the 1:1 000 000 geological map of Zimbabwe (Rhodesia Geological Survey, 1971) most of the Zimbabwean basins are massive granites of the gneissic form (e.g. EO4, EO5, EM1, EM2 and EM3 basins) while most of the Kafue River system flows on granitic forms of one description or other (Burke et al., 1994). Also underlain by granites are the South Africa basins P40 A-B and R20C (Department of Mines, 1970). The other major forms of geology in the region are the Karoo and Transvaal groups of sedimentary formations consisting of sandstone and mudstone types. For example, the lower parts of the Pungwe basin in Mozambique (Direccao Nacional de Geologia, 1983) and the Botswana basins are lowland sedimentary basins, as well as a substantial proportion of South African basins including U20D, W41A-C and V70D which exist on some derivatives of the Karoo system. At the other end of the spectrum are the basins that are underlain by one type or another of the metamorphic rock forms, e.g. the ultra-metamorphic rocks of the Sabie basin (X31A, X31A-D, X21F-K) in South Africa (Department of Mines, 1970) and the mafic or acid meta-volcanics or meta-sediments of the Mazowe River basin (D27 & D28) and the Mutare River basin (E1) in Zimbabwe (Rhodesia Geological Survey, 1971).

Some soil maps were available for all of the countries in the region. However, the scales of the maps were different and that impacted quite heavily on the interpretation of the soil texture classes. It was also necessary to have the Mozambique maps translated into English as they are produced in Portuguese. The FAO soil maps (FAO, 2003) were also accessible for the study but the scale at which they are available is too coarse and the information is thus too general at the basin scale. They were, however, valuable in providing a baseline indication of the general soil types of any given area. Notwithstanding attempts in some places to match soil unit boundaries with major landforms, the information on critical soil attributes such as soil depth and texture is not available on these maps. Data are only available for a few soil profiles and inferences could be made from these. They were thus used as a general guide and in conjunction with the available national maps. The soil maps of South Africa (SIRI, 1973), used in conjunction with the WR90 database by Midgley et al. (1994), and Zimbabwe (Department of the Surveyor-General, 1979) proved more detailed than the others with some basic soil descriptions, soil texture classes, geological source from which the soils derive and qualitative estimates of average soil depths (see e.g. Fig. 6.2). Mwelwa (2004) provides a rough guide on the climate, soils, and the geology of the Kafue basin in Zambia.



5G Mainly moderately shallow, greyish-brown, coarse-grained sands throughout the profile, to similar sandy loams over reddish-brown sandy clay loams; formed on granitic rocks.

6G Moderately deep to deep, greyish-brown, coarse-grained sands over pale loamy sands, to similar sandy loams over yellowish-red sandy clay loams or, occasionally, sandy clays; clay fraction essentially ferrallitic (no 2:1 lattice minerals), but reserves of weatherable minerals are appreciable; base saturations range from about 30 to 60%; E/C values lower than about 15 m.e.; formed on granitic rocks.

★ Location of gauging station

Figure 6.2 Illustration of the detail of soils information from Zimbabwe. The map shows part of the Mzingwane catchment with the locations of the basins for gauging stations B15, B29, B56, B77 and B78.

For South Africa, soil association (land type) maps (SIRI, 1987) at 1:250 000 scale showing units of uniform terrain form, soil pattern and macro climate were also available (Fig. 6.3). These were quite valuable in that they show detailed variations and links between the soil and other physical characteristics.

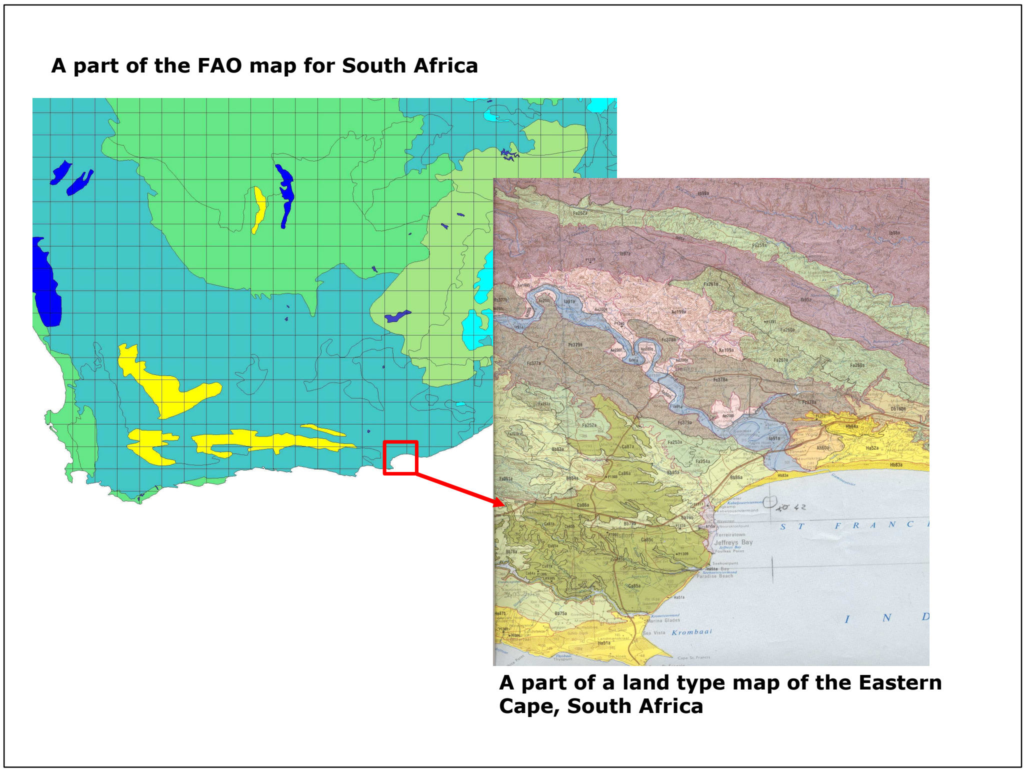


Figure 6.3 Illustration of the spatial scales and the level of detail of the soils information available in southern Africa. The figure shows a part of the FAO maps for South Africa and a land type for a part of the Eastern Cape.

The other maps had general descriptions of the soils only. The annotation of the Mozambique map was the least detailed. The availability and accessibility of information in southern Africa is therefore very variable in terms of spatial resolution and detail, from generalized FAO (2003) maps to more detailed maps such as the land type maps of South Africa (Fig. 6.3). Such disparities in the level and amount of detail available have implications for the parameter estimation methodology as uncertainties would be introduced in the areas where the detail is lacking. For South Africa, the Agrohydrological maps of Schulze et al. (2007), which became available towards the end of this project, could provide more detailed physical basin property data at the quaternary basin scale, including estimates of soil depths. The future potential of this new data source will be addressed in this section.

6.2.2 Size of basin areas

The study covered a large variety of basin sizes, with the smallest being the Mapopo River basin where flow is measured at Stapleford Forest (F1) in Zimbabwe at an area of 6.5 km^2 and the Pungwe River at Tete bridge (E66) in Mozambique being the largest at 15 046 km^2. The majority of the basins (82%) were of medium size, ranging from 100 up to 5000 km^2. In most cases, the larger the basin, the more likely it is that human influences would have altered the natural hydrology of the River system. It was therefore decided to limit the number of large size basins on the understanding that it would be difficult to access accurate data on abstractions and water use in the region. The large basins used therefore are areas where reliable quantification of the human influence has already been undertaken by previous work. Examples are the Kafue system in Zambia (Mwelwa, 2004) and the trans-boundary Pungwe system (SWECO, 2004) in Mozambique and Zimbabwe. Thus only 5 basins had areas above 5000 km^2 and all of these are made up of a number of gauged (and at times ungauged) sub-basins. For example, the Kafue at Mpatamato (gauging station number 4200), which is 11 655 km^2 in size, is made up of five upstream sub-basins of the Kafue River system and the Pungwe at Bue Maria (gauging station E66) is made up of seven upstream sub-basins of the Pungwe River system.

Such multi-sub-basin systems were important for the estimation process in that they tested its robustness. Where the intervening sub-basins had no flow records at their outlets, the model results were evaluated further downstream at the basin outlet. All the intervening sub-basins had to have their parameters quantified and it has been assumed that a representative simulation result at the basin outlet demonstrates that the values estimated for the parameters of the sub-basins were satisfactory. This is based on the further assumption that a representative result is a consequence of the ability to identify the distribution of the spatial differences in the physical properties of the sub-basins and that the estimation methods were able to convert these into appropriate parameter values. However, this assumption cannot be tested, which underlies some of the uncertainties of the estimation process in ungauged basins. It is also possible that the spatial differences were properly identified even when the simulations were satisfactory. An example is the Seekoei River in South Africa which is made up of 10 quaternary basins but with only one gauge at the outlet of the lower-most basin. While the WR90 database assumes the same parameter set for all of them, the lower-most basin has very different hydrological response characteristics from the rest and its parameters should be different. The parameter estimation methods are being developed with the eventual objective of regionalizing the model and their successful application in such large basins would indicate the potential of the principles to alleviate calibration-related problems that have previously hindered model regionalization efforts.

6.3 Applying the revised parameter estimation procedures

6.3.1 Modelling period

The overriding criterion in the choice of the modelling periods was the availability of the input data of rainfall and evaporation and the stream flow data. While the national agencies that collect these data provided some of the data, others were obtained from the International Water Management Institute (IWMI) databases (http://dw.iwmi.org/dataplatform/Links.aspx for rainfall and runoff data and www.lk.iwmi.org/WAtlas/AtlasQuery.htm for the estimates of potential evapotranspiration) while other sources such as the FRIEND database (Hughes, 1997) were also used. When the interpolated evapotranspiration data from the IWMI database were compared with the WR90 database (Midgley et al., 1994) for some South African basins, the discrepancies were minimal which increased confidence in their use in basins where the data could not be obtained from national agencies. The rainfall and runoff data for the Pungwe River basin were made available for the project from the Pungwe Basin Project (SWECO, 2004) team leader who also performed the quality checks for the data. Runoff data for the other Zimbabwean basins were obtained from the FRIEND database and the Zimbabwe National Water Authority (ZINWA), with the rainfall data coming mostly from the IWMI database. The Zambian data were available at the IWR from the work done by Mwelwa (2004) and the Botswana data from the FRIEND project database. All the River flow data for the South African basins were obtained from the Department of Water and Forestry Affairs (DWAF) website (http://www.dwaf.gov.za) and the input data were taken from the WR90 database (Midgley et al., 1994).

The modelling periods were chosen to minimize the adverse effects of missing data for most of the basins, as well as being based on the quality of the available data. The data quality analysis for the Pungwe basin was carried out during the Pungwe Basin project (SWECO, 2004) and these data were used in this study. There also was no need to revisit data quality analyses for the basins whose data were obtained from the FRIEND project, the Kafue basin and South African WR90 project (Midgley et al, 1994). There were therefore very few basins for which it was necessary to perform a preliminary data quality check. These included the Zimbabwe basins of the Macheke, Odzi, Budzi, Zonwe and Nyahodi systems. The analyses were simple and based on manual inspections of the data including visual inspection for missing data, checking negative data entries, extreme values, comparison of hydrographs with those of nearby stations, where possible, and comparison of mean annual precipitation (MAP) and mean annual runoff (MAR). The author worked in all these basins and therefore has intimate knowledge of these basins which helped in the analyses. No detailed statistical analyses were done.

The modelling periods were also chosen to avoid the influence of human activities such as afforestation, impoundments and abstractions. Of all the basins, only the Macheke River system in Zimbabwe included parameters for the Rusape dam, as a test case on human impacted basins. Mazvimavi (2003) selected basins where impoundments accounted for no more than 10% of estimated MAR and contends that a minimum of 10 years of data is required for reasonable modelling in Zimbabwe. The same criteria were used in this study. Table 6.1 shows an analysis of the distribution of modelling periods for the basins chosen in this investigation.

Table 6.1 The distribution of the lengths of modelling periods for the basins in the study

Length of period (years)	**No. of basins**
≤ 10	3
11 to 20	17
21 to 30	26
31 to 40	21
≥ 41	4
Total	71

The Pungwe River basin at Katiyo (F22) at the border between Mozambique and Zimbabwe had the shortest modelling period of only 6 years. While this is too short a period for any meaningful conclusions about the success of a model’s application, F22 is one of only two gauging stations on this important trans-boundary river basin and parameterizing this sub-basin (to simulate representative flows) was essential for the modelling of downstream basins. The majority of the basins had periods ranging from 21 to 40 years with the longest modelling period being for the Bree River at Ceres Toeken Geb (H1H003) in the H10A-C basin which was modelled for a 67 year period. Even though the start dates for South African basins were different, the end dates were always the same. All were modelled up to September 1990. The rationale for this was that it is impossible to go beyond 1990 (even though almost all the flow gauge records go beyond this date, even up to the year 2007) without having to re-process gauged rainfall data to generate new spatially averaged rain data. While this would have been possible, it would have introduced additional uncertainty given that the number of active gauges has progressively declined in the basins (see Sawunyama and Hughes, 2007). It was also considered appropriate to compare parameter values derived through the revised estimation procedures with those given in the WR90 reports by Midgley et al., (1994). Using different rainfall inputs would have precluded such a comparison.

On the whole, the periods for most of the basins were regarded adequate for modelling and water resource assessment purposes. However, it is prudent to note that, for any model, the parameter values are not independent of the climate input data. Therefore it is accepted that the quality of the input data could also have influenced the results of this study. While this was recognized in this study, time constraints could not allow an analysis of this problem. However, short modelling periods in the Mozambique part of the Pungwe basin, the basins of Botswana and the Mzingwane basins (B15, B29, B56, B77 and B78) of Zimbabwe suggest that these results should be treated with caution.

6.3.2 The revised parameters

Since South Africa is the only country with existing regionalized parameter sets, the revised parameter sets were compared with the regionalized parameter sets developed during the water resources assessment project in the 1990s, the WR90 database (Midgley et al., 1994). Table 6.2 gives a brief description of the physical attributes of a small subset of the basins investigated (see Appendix 1 for a full list of all the basins).

Table 6.2 Brief descriptions of the physical attributes of some basins investigated in the study

Country	Basin code	Gauge	Description
South Africa	K40A	K4H003	Steep topography, shallow loamy sands; fractured granite.
	G50G	G5H008	Undulating topography, moderate to deep, porous sands; unconsolidated sedimentary strata.
	V70D	V7H012	Steep topography, moderate to deep, clayey soils; interbedded mudstones, shales and sandstones.
	H10A-C	H1H003	Steep, moderately deep sandy loams; karoo shales and sandstones.
Zimbabwe	EO4	E61	Undulating topography, moderately deep sands; granites-gneissic and massive.
Mozambique	unknown	E65	Gentle to undulating topography, deep sandy clays; granites- gneissic and massive.

The parameters obtained using the revised estimation methods and the estimates of the physical property data for the same basins are given in Table 6.3. The table also includes the six objective functions used to measure the performance of the model. These are the Nash-Sutcliffe coefficient of efficiency (Nash and Sutcliffe, 1970) for both untransformed (CE) and natural logarithm transformed (CE (ln)) values, the coefficient of determination for both untransformed (R^2) and the natural logarithm transformed (R^2(ln)) values and the percentage deviation of the mean of the simulated flow from that of the observed flow for both the untransformed (%M) and the natural logarithm transformed (%M(ln)) values. The full complement of the basins, basin property data and the estimated parameters can be found in Appendix 2. It is also prudent to emphasize here that the FT, GW and POW parameters of WR90 are associated with a different version of the model than the one used in this study.

In many cases the revised parameter sets are quite different to the existing South African regional sets (Fig. 6.4). In general, the revised values of the ST parameter were almost always higher than the WR90 values (e.g. 247 against 100 for K40A), while the FT values were almost always lower (e.g. 20 mm against 30 mm for V70D). The values of the power (POW) of the moisture-interflow relationship were generally similar. Higher values of POW beyond previously expected ranges are theoretically possible in the revised procedures. The differences in runoff generated by the soil moisture function were compensated for by differences in the surface runoff parameters.

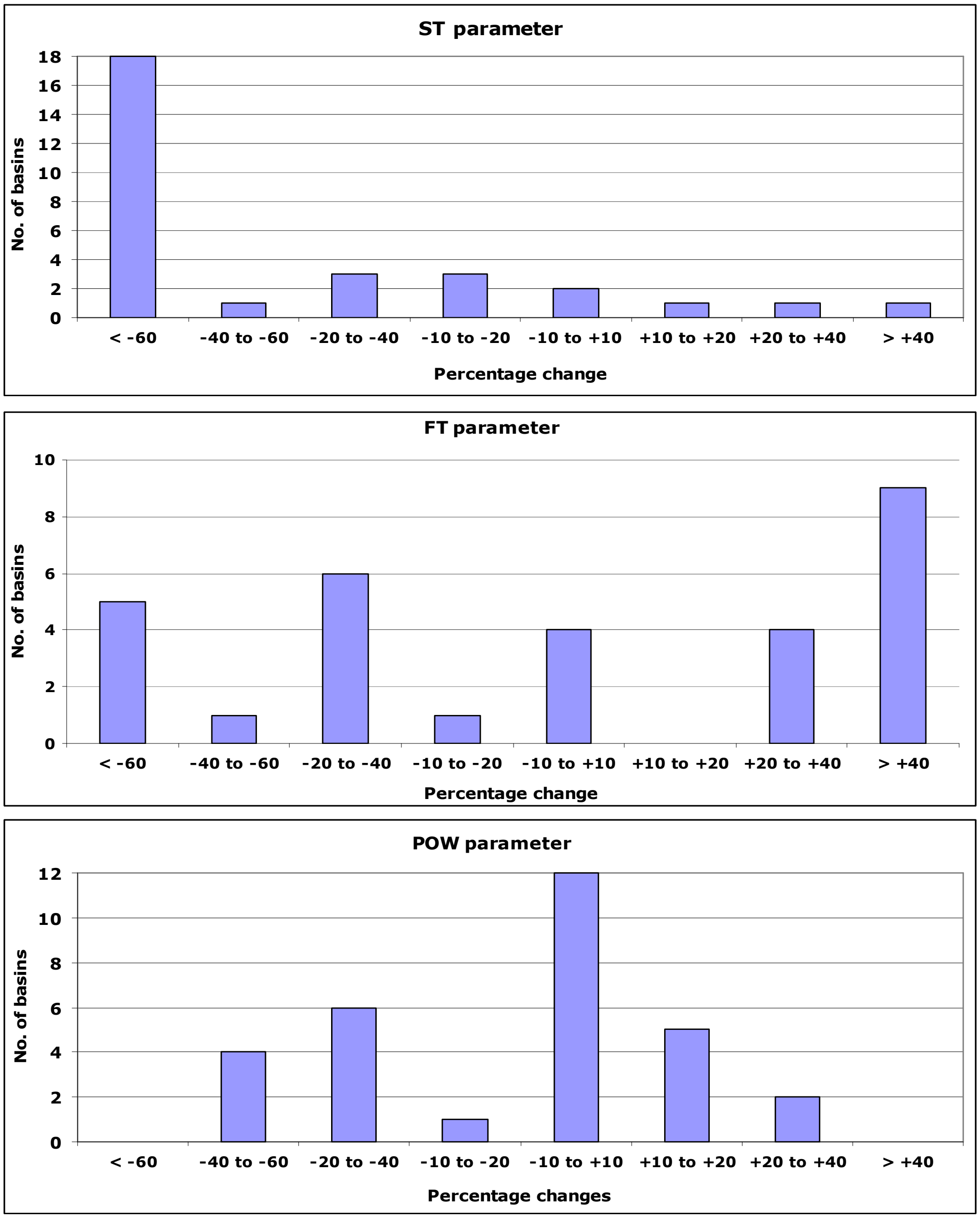


Figure 6.4 Illustration of the degree of variation of the ST, FT and POW parameter values and the number of basins with this change. The relative changes were calculated as: 100*(Revised – WR90)/WR90.

Table 6.3 Basin property data, the parameters from the physically-based parameter estimation methods and results of model simulations.

Basin	K40A	G50G	V70D	H10A_C	E61	E65
MAP (mm)	702	372	814	590	841	1092
Basin area (km^2)	72	382	196	657	2450	1313
WR90 parameters and model simulation results						
ST	100	250	120	180	-	-
FT	50	4	30	75	-	-
POW	2	2.	3	2	-	-
GW	50	5	15	15	-	-
ZMIN	0	20	999	0	-	-
ZMAX	200	350	999	450	-	-
CE / CE (ln)	0.66/0.57	0.04/0.27	0.51/0.55	0.78/0.59	-	-
Basin property data, physically-based parameters and model simulation results						
Drainage Density (km/km^2)	2.08	1.60	2.34	1.90	2.10	1.50
Mean basin slope (BS) (%)	0.30	0.08	0.30	0.30	0.12	0.15
Regional GW slope (GS) (%)	0.05	0.01	0.05	0.03	0.01	0.03
Drain. vector slope (VS) (%)	0.04	1.00	4.20	3.10	0.04	0.04
Mean soil depth (m)	0.60	0.80	0.80	1.20	1.20	1.37
FT soil depth (m)	0.73	0.84	0.93	1.53	1.38	1.69
Soil porosity	0.39	0.42	0.32	0.37	0.41	0.36
Vertical variation factor	0.80	0.44	0.80	0.62	0.80	0.80
Soil Permeability (m/day)	1.41	1.07	0.36	0.81	1.85	2.44
Depth to GW (m)	30	8	15	30	15	25
GW storativity	0.002	0.002	0.003	0.002	0.02	0.002
Unsat transmissivity (m^2/day)	5.0	2.0	1.0	5.0	2.2	2.5
ST_{soil} (mm)	187	148	205	275	394	395
ST_{unsat} (mm)	60.0	0.0	45.0	6.5	160.0	45.0
FT_{soil} (mm/month)	38.62	6.93	13.98	42.56	38.50	55.55
FT_{unsat} (mm/month)	26.21	0.00	5.90	6.50	11.64	9.45
POW	2.0	3.2	4.0	1.8	2.0	2.5
ZMIN (mm)	10	50	30	10	10	50
ZMEAN (mm)	220	100	300	110	405	350
ZMAX (mm)	250	550	550	210	800	650
CE / CE (ln)	0.66/0.67	0.36/.53	0.60/0.75	0.75/0.72	0.68/.70	0.79/0.91
R^2 / R^2 (ln)	0.63/0.70	0.63/0.55	0.62/0.75	0.75/0.82	0.71/0.70	0.81/0.92
%M	−3.7	−20.9	5.20	−4.6	−2.7	−9.7
%M (ln)	−18.8	−18.1	8.10	−24.7	2.20	−0.1

It was also a general observation of this study that where the infiltration parameters were switched off in the WR90 database (i.e. ZMIN=999 and ZMAX=999), the ST values for the basins were almost always quite small and the FT values relatively high. This ensured that sufficient runoff was generated using the moisture store routine to match the observed flows. However, this study suggests that the same surface runoff is likely to occur without switching off the infiltration parameters and in general this resulted in somewhat better model results (see V70D in Table 6.3). The same trend for the ST and POW parameter values was observed when using the shuffled complex evolution (SCE-UA, Duan et al., 1992) automatic optimization algorithm for the same model by Ndiritu (pers. comm.), though the optimized ST values were almost always greater than the revised values of this study. However, the automatic methods tend to suggest extremely high values for the FT parameter which seems at variance with the physical make up of the basin. The results from the automatic optimization suggested ZMIN values that were always very small (usually less than 10 mm). ZMAX was always higher than the WR90 values but less than the revised estimates. These observations were also true for the Kafue basin in Zambia where the automatic optimization procedure was also applied by Ndiritu (pers. comm.).

The revised parameters also seem consistent with natural characteristics; the moderately deep to deep poorly drained alluvial loamy sands of C12D sub-basin are expected to have more capacity to hold moisture than the ST value of 45 mm given in the WR90 database (Midgley et al., 1994). The revised value is 421 mm. With a less permeable subsoil that has significantly more clay from the karoo shales than the surface, the interflow is likely to be small and the value estimated by the revised method is 3 mm which compared favorably with 2 mm given in the WR90 database. With a poorly drained soil and a low gradient, it is assumed that this sub-basin would not experience rapid moisture redistribution after a rainstorm event resulting in a high value of the parameter POW, estimated by the revised method at 4.5 compared to 3 by the WR90 regionalization. This value of POW happened to be the largest estimated by the revised methodology among the South African basins investigated. The highest values of ZMAX were estimated for the areas of more permeable soils and, often but not always, low slopes (e.g. 800 mm for sub-basins X31A-D). The lowest value of ZMIN was 0 mm which was associated mostly with areas of surface clays that tend to crust during the dry season and would significantly reduce infiltration and therefore allow surface runoff generation to develop even from low rainfall amounts and especially at the start of the wet season (e.g. the sub-basins W52A-C). In spite of the differences in the parameter sets, the correspondence between the simulated and the observed flows indicated that the parameters of the revised physically-based estimation methods produced results that were at least as good as the current regionalized parameter sets, and in most cases even better (Fig 6.5 and Tables 6.4 and 6.5). While the results of the simulations are encouraging, there were some notable differences, in some situations (e.g. basin V70D, Fig. 6.5A), between modelled and observed hydrographs. These could be attributable to model structural weakness, poor quality observed streamflow data where human influence on the natural hydrology of the basin is usually ill-defined, or to inadequately representative rainfall input data. It is very difficult to determine which of these possible influences will dominate in any specific basin.

In the other parts of the region where no regional parameters exist, model simulations were compared with observed flows only. The parameter values estimated through the revised estimation methods were consistent with natural phenomena and therefore physically plausible. The highest ST_{soil} value was 1080 mm for the low lying, deep weathered granitic sandy loams of the Mutare River basin (gauging station E1) while the lowest was 227 mm for the shallow sandy soils of the Lumani River system (gauging station B15), both in Zimbabwe. The highest value for ST_{unsat} was 1040 mm which was estimated for the fractured schists, gneisses and granulites of the Luswishi River system (gauging station 4340) in Zambia. The highest value for the ST parameter was 1638 mm for the Zonwe River sub-basin measured at gauging station F10 and the lowest was 268 mm for the Lumani River system (gauging station B15). Both sub-basins are in Zimbabwe. The former is characterized by a deep deep-weathered mantle underlain by quite heavily fractured granite rocks and the latter is an arid basin of shallow sandy loams.

FT, the runoff from the subsurface storage when the basin is saturated, was quantified by estimating the outflow from both the soil and unsaturated components. The highest outflow value from the soil component was 118 mm which was estimated for the steep, deeply weathered granitic sandy loams of the Pungwe River basin (gauging station F14) in Zimbabwe, while the lowest was 0 mm for the shallow sandy soils of the Metsomethlaba River system (gauging station 2411) in Botswana. The same two basins had the highest and lowest outflows from the unsaturated zone (FT_{unsat}) at 33 mm (F14) and 0 mm (2411) respectively. The highest overall outflow values (FT) from the subsurface moisture zone were estimated for steep, well vegetated, deep weathered sandy loams (with some clay lenses) sub-basins of the Pungwe River system in Zimbabwe (150 mm, gauging stations F14 and F22) and the Mwambashi sub-basin (100 mm, gauging station 4120), a tributary of the Kafue River system in Zambia. Both basins are humid with rainfall totals of at least 1500 mm per annum.

A. Flow simulations using WR90 and the revised parameters at V7H012 (V70D), Little Boesmans River, South Africa

B. Flow simulations using WR90 and revised parameters at H1H003 (H10A-C), Bree River, South Africa

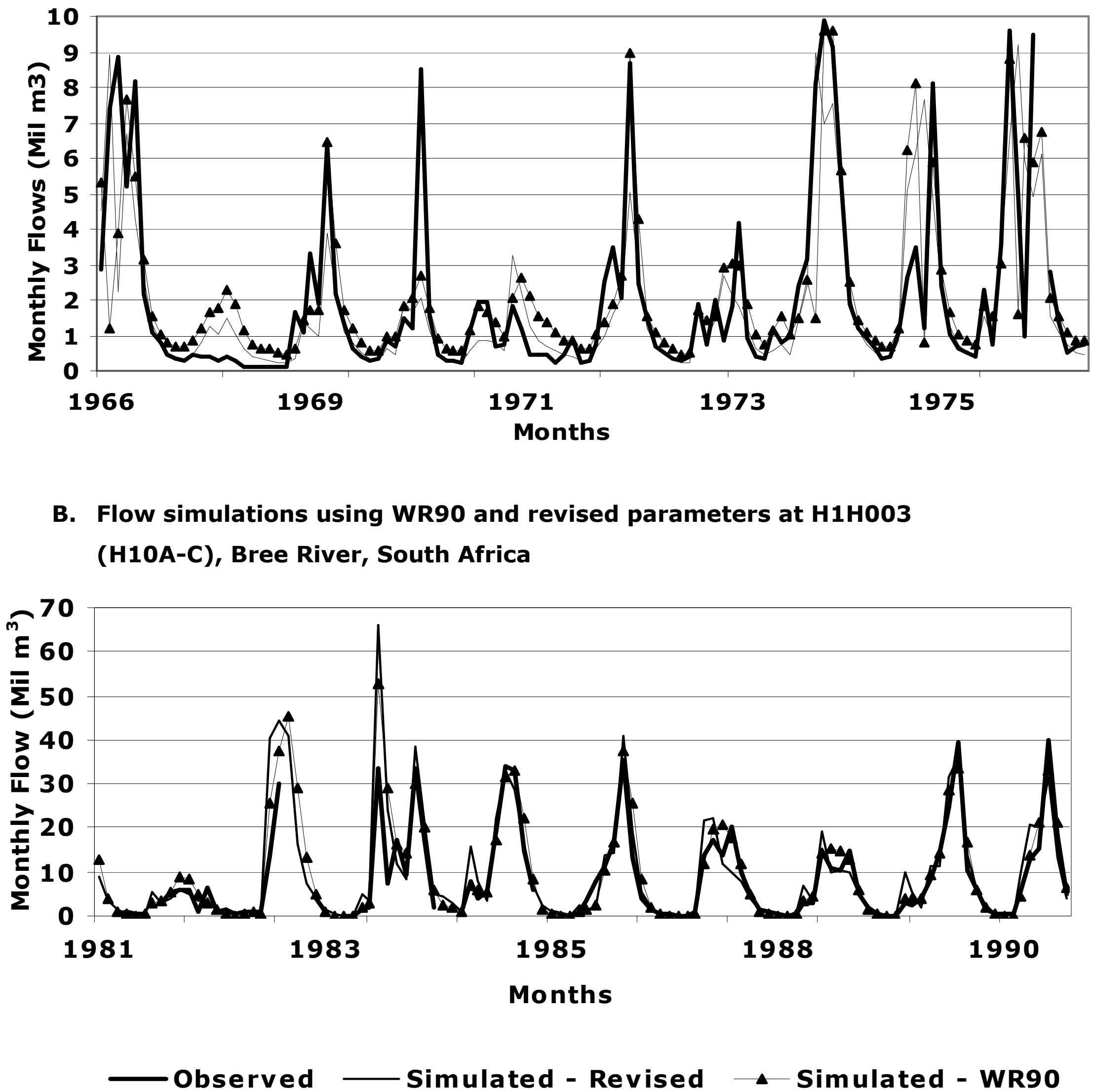


Figure 6.5 Results of model simulations using WR90 and the revised parameters compared to the observed flow for the Little Boesmans River at gauging station V7H012 (A) and the Bree River at gauging station H1H003 (B).

The Pungwe sub-basin (F14) also had the lowest estimated value for the POW parameter of 1, based on steep slopes and well drained soils. The lowest values for the overall FT parameter were estimated at 1 mm for the Botswana basin of Metsomethlaba at gauging stations 2411 and 2421 and 4 mm for the rocky (and thin soils) sub-basin of the Macheke River (gauging station E19) in Zimbabwe.

The highest value of the parameter POW was estimated at 5 for the Mvumvumvu sub-basin (gauging station E125) in Zimbabwe, the Pungwe sub-basin (in Mozambique) at gauging station E66 and the Luwishi sub-basin (gauging station 4340), a tributary of the Kafue in Zambia. These sub-basins are characterized by almost homogeneous soil distributions which ensure uniform soil wetness conditions in generally very gently sloping to undulating landscapes. This ensures a slow redistribution of moisture after rainstorm events.

The lowest value of the parameter ZMIN was zero and the highest value of ZMAX was 1450 mm for the Mvumvumvu sub-basin in Zimbabwe and the Luwishi sub-basin in Zambia also had a high ZMAX value of 1400 mm.

The results of the modelling simulations using the parameters estimated by the revised methods, albeit with some calibration of the other free parameters, gave good results (e.g. Fig.6.6). The results indicated that revised estimates of the moisture accounting and main runoff producing parameters were able to reproduce the main characteristics of the hydrology of the selected basins. The results show that the means of the simulated flow time series for almost all the basins were within +/-10% of the means of the observed flow time series and the other objective functions indicated that the model was able to explain at least 60% of the time series variation of observed flows.

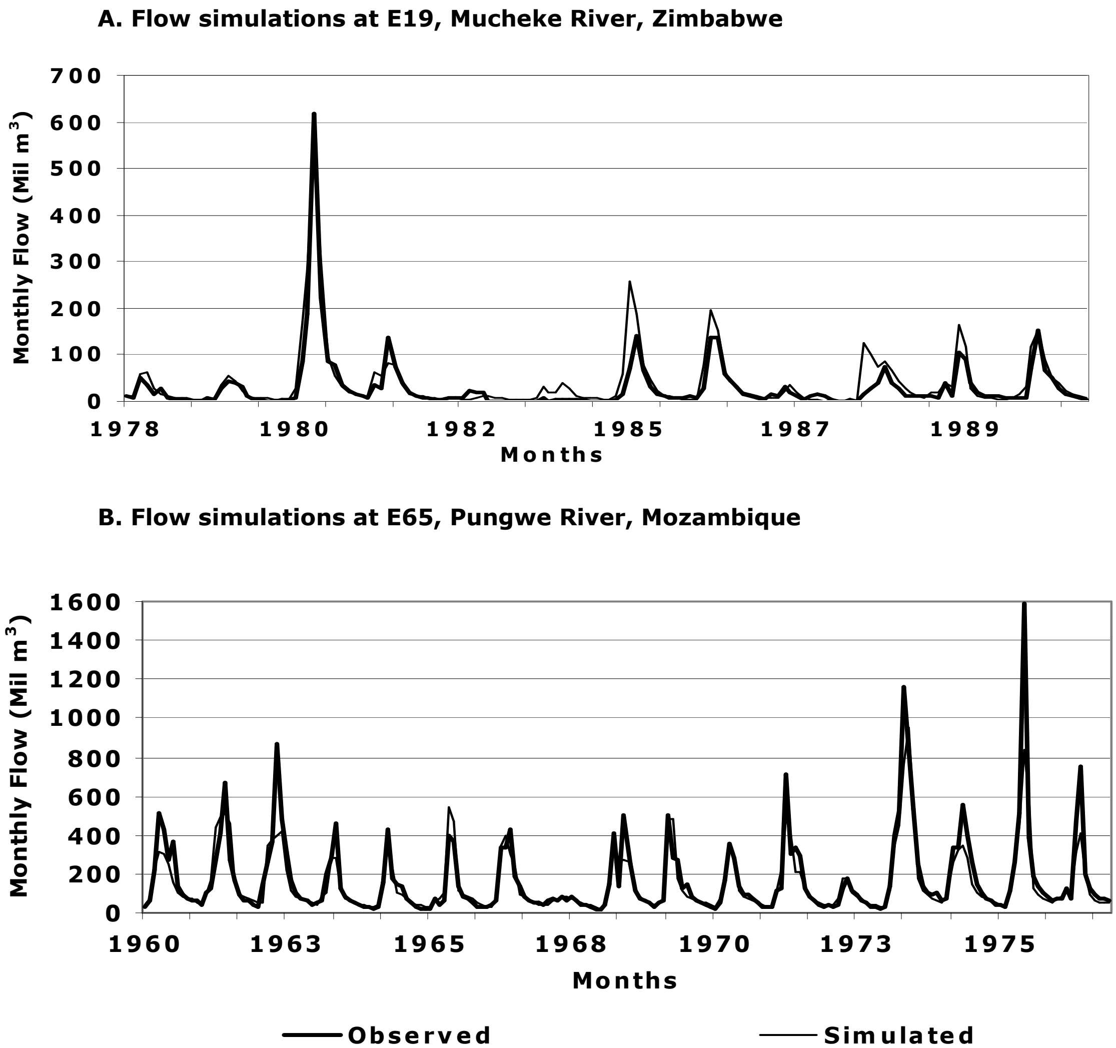


Figure 6.6 Results of simulations using the revised parameter estimation methodology for the Macheke system at gauging station E19 in Zimbabwe (A) and for the Pungwe River system at the Pungwe Bridge (gauging station E65) in Mozambique (B).

6.3.3 Measures of model Performance

Six objective functions were used in this study to assess the performance of the model. For the South African basins the objective functions effectively gauged the appropriateness of the revised parameters since the model was not calibrated against observed flows. The GW parameter had to be adjusted to generate assumed recharge rates (based on the Groundwater Resource Assessment II project database - DWAF, 2005) after the other parameters were changed. The values of the parameters that were not part of the revised estimation scheme were kept the same as the original WR90 parameters. The analysis of four of the objective functions (CE, CE (ln), R^2 and R^2 (ln)) is given in Table 6.4, while the percentage errors (%M) are compared in Table 6.5.

Table 6.4 An analysis of the results of model simulations based on the four of the six objective functions used.

Range of objective functions (%)	**CE**			**CE (ln)**			**R^2**			**R^2 (ln)**		
	A	B1	B2	A	B1	B2	A	B1	B2	A	B1	B2
› 90	0	1	0	1	0	0	0	1	0	1	0	0
80 – 89	3	2	2	9	6	2	5	3	4	10	8	7
70 – 79	16	8	3	8	10	6	20	8	8	7	12	9
60 – 69	16	8	11	10	9	6	11	11	9	11	6	7
≥ 60 subtotal	**35**	**19**	**16**	**28**	**25**	**14**	**36**	**23**	**21**	**29**	**26**	**23**
50 – 59	4	7	3	8	5	7	4	5	3	8	4	5
40 – 49	1	3	1	1	0	3	0	2	5	1	0	2
‹ 40	1	1	10	4	0	6	1	0	1	3	0	0
< 60 subtotal	**6**	**11**	**14**	**13**	**5**	**16**	**5**	**7**	**9**	**12**	**4**	**7**

Note that Column A represents the simulation results for the other southern African basins, B1 is for South Africa basins using the revised estimation method and B2 is for South African basins using WR90 regionalized parameter sets.

In this study the standard for a successful simulation for any given basin, and therefore successful parameter estimation, was set to an objective function value of at least 0.60, i.e. if the model is able to explain at least 60% of the observed time series variation or that the synthetic indicator of the internal efficiency of the model (and its parameters) is at least 60%. This figure is assumed reasonable given the unknown and unquantified uncertainties in the model input data and the observed flow. For instance, in most cases in the region the high flows often overtop the capacity of the measuring instruments and these are estimated using extrapolation equations developed for the particular measurement point. This therefore introduces some uncertainty in the recorded high flows. The seemingly unrealistic high flow recorded at F14 on the Pungwe River in Zimbabwe for the month of March 1976 dwarfs all the measurements around it, including other high flows, even though the rainfall record does not seem to support the occurrence of such a flood (Fig. 6.7). While this high flow value may be uncertain, an analysis of the historical rainfall records in the area suggests that it can not be discarded (SWECO, 2004).

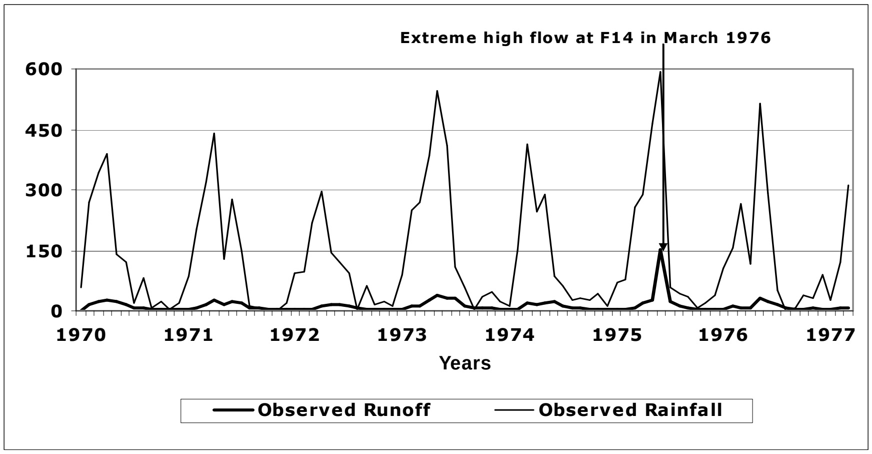


Figure 6.7 An uncertain extreme high flow value in the Pungwe basin at F14 recorded for March 1976.

On the other hand, the progressive shrinking of the precipitation gauging network also introduces uncertainties in the simulation results (Sawunyama and Hughes, 2007). Thus, given the other undefined uncertainties in the quality of the input data and the observed flow data it is often difficult to get very high values for the objective functions. Therefore the CE and R^2 values of 0.6 could be the best that may be reasonably expected and such values are usually acceptable in water resources assessment studies in the region. On this basis, the CE results reflect that the revised parameters were successful in 85.4% of basins chosen from the other southern African countries without regionalized parameter sets and 63.3% of the South African basins. The same South African basins had 53.3% successful simulation CE results when the WR90 parameters were used, which meant that there were more basins meeting the goodness-of-fit criterion when using the revised estimation methods than with WR90 regionalized parameter sets.

The trend is the same when using the other three objective functions as seen in Table 6.4 with the percentage of successful simulations for objectives functions CE(ln), R^2 and R^2(ln) being 68.3%, 87.8% and 70.7% respectively for the other parts of the region outside South Africa. For the South African basins, when using the revised parameter sets, the percentages are 83.3%, 76.7% and 86.7% while using the WR90 parameter sets the percentages are 46.7%, 70.0% and 76.7% respectively. These results suggest a definite improvement in the low flow simulations which may be partly a result of the revised model (Hughes, 2004a) and partly related to the parameter estimation procedures. It is quite difficult to separate out these two influences.

The last two objective functions measured the percentage deviation of the simulated mean annual runoff from that of the observed record for both the untransformed (%M) and the natural logarithm transformed (%M (ln)) values. The standard for a successful simulation was set at a deviation within +/-10%. The rationale for choosing this is the same as for the other objective functions and that this constraint is adequate for the capture of a system's major flow regime without deviating far from the observed. Table 6.5 shows an analysis of the results of the simulations as measured by the %M objective function.

Table 6.5 Analysis of model simulations using the percentage deviation of the simulated mean flow from the mean of the observed flow (%M).

Range of objective functions (%)	**%M**			**%M (ln)**		
	A	B1	B2	A	B1	B2
≤ 5.0	19	17	3	21	8	2
5.1 – 10.0	20	4	5	14	7	8
≤ 10 subtotal	**39**	**21**	**8**	**35**	**15**	**10**
10.1 – 15.0	1	4	5	3	4	2
15.1 – 20.0	0	3	5	0	6	1
20.1 – 25.0	0	2	2	1	2	2
≥ 25.1	1	0	10	2	3	15
> 10 subtotal	**2**	**9**	**22**	**6**	**15**	**20**

Note that the deviations can be negative (i.e. underestimation of the mean of observed flow) or positive (i.e. over-estimation), only the magnitudes are represented in the table. The notation for the grouping of the basins (A, B1 and B2) is as defined for Table 6.4

For South African basins, it is evident from Table 6.5 that the simulations using the revised estimation procedures fared much better than those using the WR90 regionalized parameter sets – 70.0% against 26.7% for %M and 50.0% against 33.3% for %M (ln). For the basins outside South Africa the model performed equally well with 95.1% and 85.4% with deviations within +/- 10% of the observed flows for %M and %M (ln) respectively.

Overall, in spite of the considerable differences in the physiographic and climatic conditions of the 71 basins in which the methodology was applied, the revised parameters managed to account well for these differences and gave satisfactory and encouraging simulation results.

6.3.4 Revised estimates of basin physical properties

Towards the end of the project an additional data resource became available for South Africa. This was in the form of the South Africa Atlas of Climatology and Agrohydrology (hereinafter referred to as the Atlas) by Schulze et al. (2007) which has descriptions of physical basin property data for all basins in South Africa. The Atlas contains detailed maps of climatic variables and estimates of some physical basin attributes, at a regional scale, and is intended "to provide the 'big picture' in South Africa, but in sufficient detail to be useful in regional and local decision making" by practitioners in the water and agriculture sectors. In order to achieve the regional perspective the approach used was based on extrapolations of the gauge records to ungauged places. From the Atlas one is able to access climatic (e.g. rainfall, temperature, solar radiation, vapour pressure), hydrologic (mean annual runoff, baseflow, mean monthly sub-basin flows, etc), soil (e.g. hydraulic conductivity, porosity, depth, etc) and agricultural crops data. These data are very valuable for a study of this type as they have the potential to provide the relevant basin attributes data. Of particular interest were the estimates of soil property data and especially soil depth. The project had so far relied on descriptive estimates of soil depth contained in the WR90 report (Midgley et al., 1994) for guidance. It was anticipated that the availability of the Atlas would provide improved quantification of this attribute. The depth estimates from the Atlas were derived from the SIRI (1987) land type maps which have quantitative typical soil profile total depth range estimates. The Atlas divides the total soil profile into the topsoil and the subsoil and the depths for each of these are estimated from the land type maps (Schulze, 2006). These two components of the total soil profile are assumed to coincide with this study's conceptual 'soil' component. While the SIRI (1987) land type maps provide for the soil depth to be greater than 1.2m, the total depth in the Atlas has a maximum of 1.5m (Schulze, 2006). However, depths greater than this are possible in some basins in South Africa and in the WR90 qualitative guidelines it is assumed that soils deeper than 2.0m exist (Midgley et al., 1994). The Atlas also has estimates of soil porosity which are relevant to the parameter estimation methods of this study.

Table 6.6 shows the soil depth and porosity data obtained from the Atlas compared to the estimates initially used in the parameter estimation methods. In general the mean basin soil depths estimated in the Atlas are lower than the estimates used in the estimation procedures whose values were guided by the qualitative descriptions of the Midgley et al. (1994) and the SIRI (1973) soil map. However, most of the soil depths in the Atlas were within the qualitative descriptions and this, in a way, demonstrates the uncertainty associated with the use of the ranges covered by the descriptions. For example a 'moderate to deep' category is from 1.0m to 2.0m. The porosity values in the Atlas are higher than the estimations from the USDA (1969), Cosby et al. (1984) and Rawls et al. (1982) which were used to guide the values used in this study.

Table 6.6 Comparison of estimated soil depth and porosity used to assess data from the Atlas against the estimates used in this study.

	Porosity		**Soil depth**		**Texture class**	**WR90 depth**
	Estimate	**Atlas**	**Estimate**	**Atlas**		
G10A-C	0.42	0.49	1.00	0.51	Sandy loam	Shallow to moderate
X31A-D	0.39	0.47	1.60	1.03	Clayey loam	Moderate to deep
W52A-C	0.35	0.49	1.00	0.87	Sandy loam	Moderate to deep
V70B	0.35	0.47	0.80	0.68	Clayey	Moderate to deep
X31A	0.38	0.47	1.00	1.01	Clayey	Moderate to deep
G10B	0.40	0.45	1.00	0.40	Sandy	Moderate to deep
K60A	0.40	0.49	0.36	0.37	Sandy loam	Shallow to moderate
V20A	0.39	0.47	1.20	0.81	Clayey	Moderate to deep
P40A-B	0.38	0.47	1.20	0.61	Clayey loam	Moderate to deep
K20A	0.38	0.47	1.20	0.59	Clayey loam	Moderate to deep
C12D	0.35	0.45	0.67	1.50	Sandy	Moderate to deep

Depth guide: < 0.2 ≈ shallow; 0.2-0.5 ≈ shallow to moderate shallow; 0.5-1.0 ≈moderate deep; 1.0-2.0≈ moderate deep to deep and >2.0 ≈ deep

Eleven basins were chosen to test the data from the Atlas in the parameter estimation procedures. Only the soil depth and porosity data were used. The rationale was that these data were regarded as crucial in the estimation procedures and the values that had been used were quite uncertain and need improving. The soil depth and effective porosity values were substituted into the relevant algorithms. The depths of the topsoil and subsoil were added to make up the 'soil' component. In the absence of any data on the unsaturated component of the soil profile, the estimated physical property data used in the initial estimations were adopted. The resulting parameters were then used in the model, without recalibration, and their suitability assessed. A summary of the estimated parameters and the results of model simulations in four of the chosen test basins are given in Table 6.7. The maximum soil moisture content parameter (ST) is almost always lower than the initial estimates. This is explained by the lower soil depth values from the Atlas. In general, this means that the amount of interflow at saturation (FT) from such restricted depths is likely to be smaller as well. The POW parameter was generally similar throughout all the basins used.

While a decrease in the value of the parameters ST and FT implies that within the model a lower maximum infiltration rate parameter (ZMAX) is required in order to compensate and simulate the high flows correctly, it seems that this did not work here. Except for the Berg basin (G10A-C), whose results were almost similar for the different data sets, the results indicate the Atlas data gave inferior simulation results. Even for the headwater sub-basin of the Sabie (X31A) where the Atlas depth of 1.01m was very close to the initial estimate of 1.0m, the simulation results are very different owing to different effective porosity values. The Atlas gave a porosity of 47% for the sandy clayey loam soils against the initial estimate of 39%. The use of the Atlas values offers a more objective approach to estimating some of the physical basin properties in South Africa. However, they do not appear to be consistent with the Pitman model parameter estimation procedures that have been developed in this study. This issue requires further investigation if the Atlas values are to be considered for use in estimating the parameters of the Pitman model.

Table 6.7 Estimated parameters and the results of model simulations for 6 of the basins using the basin property data from the Atlas (Atlas data) compared with the initial parameter estimates using the revised procedures.

Basin	**Parameter**	**Revised**	**Atlas data**		**Basin**	**Parameter**	**Revised**	**Atlas data**
G10A-C	ST	294+30	250+24		W52A-C	ST	210+20	256+19
	FT	22.4+17	19.1+4.8			FT	8.1+0.4	12.7+10.1
	POW	2.2	2			POW	2.5	2
	GW	30	12			GW	15	8
	ZMIN	50	30			ZMIN	0	50
	ZMAX	400	520			ZMAX	325	500
	CE/CE(ln)	0.89/0.83	0.88/0.83			CE/CE(ln)	0.41/0.66	0.23/0.49
X31A-D	ST	437+47	238+50		V70B	ST	224+40	256+50
	FT	47.5+17.1	6.6+6.2			FT	16.1+2	3.2+2.8
	POW	2.80	2.5			POW	2	2
	GW	50	35			GW	20	20
	ZMIN	0	0			ZMIN	10	0
	ZMAX	800	500			ZMAX	440	250
	CE/CE(ln)	0.82/0.82	0.65/0.72			CE/CE(ln)	0.67/0.77	0.48/0.49
X31A	ST	274+20	342+20		G10B	ST	320+40	250+24
	FT	11.2+6.8	6.3+6.8			FT	20.1+14.1	19.1+4.8
	POW	2	2			POW	2	2
	GW	60	50			GW	5.5	2.5
	ZMIN	0	0			ZMIN	20	0
	ZMAX	750	480			ZMAX	380	400
	CE/CE(ln)	0.74/0.71	0.53/0.66			CE/CE(ln)	0.74/0.84	0.61/0.85

7 DISCUSSION AND CONCLUSIONS

7.1 The parameter quantification approach

One of the critical issues in this study has been the need to assess the effect of each model parameter on the basin scale runoff processes and investigate their physical meaning. This is an essential component of the development of conceptual relationships between the parameters of a model and physical basin characteristics. To achieve this, it is necessary that the model be conceptually sound and able to adequately represent the more significant basin hydrological processes. Further, the model should not have parameters that represent the effects of multiple processes, but the fewer the parameters the more difficult it is to avoid this problem. This represents a deviation from the often stated principle that models should be parsimonious and have as small a parameter space as possible. However, the quest for parsimony is associated with the desire to achieve identifiable parameters and avoid equifinality, while the approach to resolving these issues in this study is different. Adequate process representation and parameterization at the appropriate scale enable the separation of processes and therefore isolation of the parameters of the model for examination. The assessment and separation of the effects of individual parameters facilitates their conceptual hydrological interpretation. It is then possible to identify the physical basin attributes that will be appropriate for the development of possible relationships. The Pitman model is a conceptual, well structured model built on sound hydrology principles and represents individual components of the runoff process at the basin scale on a monthly basis. It has thus been possible to develop a physically-based conceptual framework for the estimation of its parameters using physical basin attributes. This has been achieved through the re-interpretation of the physical meaning of the parameters of the model.

The soil moisture store, runoff, recharge and infiltration parameters have been quite successfully physically defined in this study and hydrologically sensible relationships have been developed. For example, on the basis of the model structure, it is reasonable to assume that the maximum soil moisture store (ST) can be divided into two components (the soil and unsaturated components). The amount of moisture that the soil component holds would logically depend on the soil's porosity and its depth and the unsaturated zone capacity would be influenced by the storativity and depth of the fractured zone. In general, it was thus postulated that deep, well-drained soils and gentle slopes have the capacity to hold more water resulting in a higher value for ST while shallower soils, often more characteristic of steeper headwater basins, have lower ST values. It was also assumed that the release (rate and magnitude) of the moisture stored in the basin as interflow (both rapid and delayed) would depend on the drainage density, surface topography and the saturated hydraulic conductivity, as well as the ability of the underlying geology to transmit the moisture from the unsaturated zone through fissures, cracks, crevices or through perched aquifers related to impermeable lenses or layers. Variations in interflow rates are expected to depend on the volume of moisture available in the basin and its spatial distribution. This spatial distribution was assumed to be a function of basin slope and soil drainage properties which determine the rates and patterns of moisture re-distribution after storm events.

To quantify POW it was necessary to understand the process of interflow and how it is influenced by the moisture distribution of the basin following a rainstorm event. It was envisaged that the low-lying areas would naturally stay wetter, and therefore contribute interflow for longer periods, than the steeper areas (the partial and variable source area hydrological concepts). Thus the ability of the basin to redistribute moisture has a critical influence in determining the shape of the moisture-interflow relationship. The assumption was made that, within a basin, moisture would move slowly in gently sloping areas with poorly drained soils, whereas in steeper landscapes with well-drained soils it would be quickly re-distributed. These extremes can be interpreted to give different values of the POW parameter. Basin slope, soil type and characteristics (which determine its ability to transmit moisture), as well as the underlying geological formation were deemed to be the dominant physical factors.

The infiltration parameters are essentially a function of the soil surface conditions, the size of the soil moisture store, number and spacing of rain days (which influence the antecedent moisture conditions at the start of a rainstorm event) and typical storm durations (indicative of expected rainfall intensities). The Pitman model does not explicitly include the effect of saturation excess type runoff in the infiltration surface runoff generation algorithm. However, the development of an appropriate estimation process found that this process could be important and therefore included the maximum soil moisture store as a factor.

After the conceptualization of the processes and the identification of the relevant physical basin properties that influence the different parameters, relationships were developed based on physical hydrology principles. Generally, the relationships are considered adequate and appear to produce hydrologically sensible parameter values. The estimated parameter values have also been demonstrated to adequately simulate the hydrology of the selected basins. This implies that the conceptual framework and the resultant relationships are credible.

Two issues warrant further attention. The first is the availability of physical basin property data within the southern Africa region and the differences in the level of detail of the data that are available. Of great concern is the absence of specific information on soil depths in all the soils data available. Such data paucity introduces uncertainty in the estimation procedures. This problem will require further attention if the procedures are to be valuable for the more widespread regional application of the model.

Secondly, the model operates at the basin scale and at a monthly time step, while it is acknowledged that hydrological processes typically manifest at smaller scales. Within this study frequency distributions have been used in an attempt to resolve some of these scale issues. The procedures for estimating the parameters related to the infiltration process (ZMIN, ZMAX) and the power (POW) of the moisture-interflow relationship have relied on the use of frequency distributions and the probability distributed principle of Moore (1985). While this has been reasonably successful, there remain other scale issues that are difficult to resolve. These are related to the integration of basin-wide variations of such properties as soil depth, slope, hydraulic conductivity, transmissivity, etc. Without established integration methodologies, it was not possible to avoid an element of subjectivity in the basin average values that were used and this introduces further uncertainties in the estimation process. Further work is therefore required to develop relationships that can provide more objective basin average values using point or small scale observations. The methods used in this study have produced generally acceptable results, but improvements will rely on the further resolution of scale issues and the availability of more appropriate basin property data.

7.2 Evaluation of simulations using the revised parameters

The representativeness of model simulations is generally influenced by the quality of the input data and with some data sets it is not possible to get satisfactory simulation results at all. In this study it has been recognized that the quality of the input climate data may have played a role in the quality of the simulations. There are also uncertainties related to the extent to which the available observed flows represent the natural hydrology of the basins. Human influences on most rivers within the region in the form of small scale river (and off river) storages (farm dams), return flows and run-of-river abstractions are inadequately quantified.

Under all the physical and climatic conditions used for this study, the quality of the model simulations is encouraging. It is, however, possible that any inconsistencies in the parameter estimation process may have been offset by the calibration of the other parameters in the model, especially in the basins outside South Africa where no regionalized parameter sets exist. The only way to overcome this potential problem of 'compensatory calibration' would be to include all calibration parameters in the estimation procedures; therefore, future work needs to address this issue. However, the values of all the calibrated parameters are physically plausible and within expected ranges. The success of the revised simulations in the South African basins, where none of the other parameters were changed, tends to suggest that this is not really a problem. In the case of South Africa, while the estimated parameters were often quite different from the WR90 regionalized parameters, which have become 'conventional wisdom', the revised simulation results were similar and frequently better. Outside South Africa the results were equally good, suggesting that the estimation procedure is quite robust. This suggests that the Pitman model is a conceptually realistic model with conceptually realistic parameters that can be broadly interpreted (and quantified) using physical hydrology principles and measurable basin-scale physical attributes.

It is also prudent to highlight here that the finding that automatic calibrations indicate the same trend of parameter values though substantially higher FT values than the revised estimates has important implications. While a comprehensive comparison of results of this study with those obtained using automatic calibration (or Monte Carlo) approaches would be valuable, further information about the 'real' active processes would be required to determine which set of parameters leads to the most 'behavioural' (Beven, 2001) model.

7.3 Evaluation of physical basin attributes data in the region

In spite of the relatively encouraging results, there are a number of sources of potential uncertainty with regards the estimation methods. Firstly, within the southern African region the appropriate physical basin data are not easily available. Where the data are currently available there are considerable disparities in the spatial scales and the levels of detail. One of the motivations for this study is that with the developments in GIS and remote sensing technologies, these data may soon be more widely available and the model parameter estimation procedures may need to develop accordingly. Currently there are uncertainties associated with the estimation of appropriate basin-scale soil texture class, depth, hydraulic conductivity, transmissivity and storativity values. While fairly reliable point, or small scale, estimates of most of these properties could be available in the literature (e.g. Rawls et al., 1982; Schulze et al., 1985) their direct use may not be appropriate given the differences between the scales of data acquisition and modelling. Therefore methods of extrapolating from point data to the basin scale may need to be developed. This would greatly enhance the reliability and objectivity of the estimation procedure. It is also suggested that the widespread application of this revised approach may prompt improvements in the acquisition of the relevant data resulting in improved availability and accessibility within the region.

While the recently published South African Atlas of Climatology and Agrohydrology (Atlas) (Schulze et al., 2007) and SIRI (1987) land type maps were not fully evaluated in this study, their limited application raised some issues that need highlighting. Though the land type maps are a good source of information on soil properties, they are largely qualitative, while hydrological modelling demands the quantification of the physical attributes that control basin hydrological processes. This problem limits the use of the land type maps as the information that they contain is not directly relevant to a study of this kind and some form of information reprocessing is necessary. There are also cost issues related to their acquisition.

The first impressions of the Atlas information on soil hydraulic properties were favourable. The Atlas represents a recent advance in producing the type of data required in a format and spatial resolution that is suitable for this study. While the limited number of test basins makes it difficult to reach firm conclusions, it is unfortunate that some of the data appear to be inappropriate. The data on soil depth and porosity extracted from the Atlas information were quite different from other estimates made as part of this study. The main problem appears to be a generally subjective upper limit imposed on the Atlas lower horizon depths. There seems to be little doubt that deeper soils can be found in many basins of South Africa, even when averaged over the mapping scale of the Atlas. The Atlas depth estimates of the two soil layers may be appropriate for agricultural use and with agrohydrological models with a number of soil layers (such as ACRU). However, they do not seem to be appropriate for use with conceptual, semi-distributed hydrological models such as the Pitman model. Further investigations are required to assess, for example, why there are such low upper limits on the lower horizon soil depths. This may be related to its agricultural use but further clarity is required. This highlights the fact that rarely are data on physical basin properties produced for hydrological purposes and generally have to be re-interpreted or re-processed in some way. While it is difficult to make any conclusions based on only eleven test basins, it seems reasonable to assume that the direct use of the data from the Atlas, in its current form, would not generally improve modelling results for the Pitman model.

The availability of data in the non-South African countries is likely to remain a problem in the immediate to medium-term future. While the use of GIS and remote sensing technologies may improve the extent and quality of appropriate physical property data, it may be some time before most of the countries in the region start using them. The technology may be beyond the means of most of the countries in the region that have other more pressing priorities. The shrinking measurement networks for hydro-meteorological variables illustrate the inability of the majority of countries in the region to maintain important data collection platforms due to constrained financial resources. Despite this rather negative outlook for the immediate future, there is little doubt that the incorporation of remotely sensed and GIS data would make physical basin property data more readily available and accessible especially in areas of scarce input data.

7.4 Conclusions and Recommendations

This study has established the potential of the use of physical basin characteristics in the direct quantification of the parameters of the Pitman model, which has important implications for the regional application of the model. In general the revised parameter estimation methods have generated physically reasonable values which have been used to satisfactorily simulate the hydrology of the basins investigated in this study, when compared to previous simulations.

The main focus of this study has been on developing a conceptual framework and establishing linkages between parameters and the physical basin properties. The results contained in this report therefore represent preliminary tests. The parameter estimation procedures have been tested using limited, coarse spatial resolution information. However, the encouraging results suggest that they are fairly robust and conceptually correct. It has been shown in this study that it is quite possible to isolate the effects of each parameter and develop relationships with physical basin attributes, implying that the parameters of the model are physically meaningful. The equations developed for the estimation procedures are based on physical hydrology principles and have been demonstrated to produce satisfactory simulation results. However, there remain a number of uncertainties which are partly related to the quality and resolution of the physical basin data. As more relevant data become available further tests would be required and it is possible that revisions to the estimation procedures may also be required.

For the South African basins, the simulation results have been compared to previous results using regionalized parameter estimates and an earlier version of the model (as well as observed data). It has been concluded that the parameter estimation procedures that have been developed have contributed to the improved results. However, it is also reasonable to suggest that the revision of the model (inclusion of more explicit surface-ground water interaction routines) also played a significant role. It is difficult to separate the influence of the new version of the model from that of the estimation methods, as the estimation methods are developed for the new version. However, previous work on the revised model, but using similar regional parameters as used in WR90, indicated that the simulations are very similar to the WR90 simulations (Hughes and Parsons, 2005). Thus, while it is not possible to entirely separate the influences, the indications are that the parameter estimation procedures play the major role in the improvements.

Given the diversity of the physical characteristics of the basins used it would be safe to conclude that the estimation methods are robust. The methods should contribute to more consistent and objective parameter quantification and improve the potential to apply the model in ungauged basins without reliance on calibration results. One of the motivations for this study was the eventual development of regional parameter sets similar to those established during the South African water resources assessment project of the 1990s (Midgley et al., 1994). While this was not possible during the limited time available in the present study, the development of calibration-free parameters of this nature is a first step in the right direction. The methods should contribute to the regional application of the model which is a practical requirement of water resource managers who are often called upon to make hydrological predictions in data scarce areas for long term, often highly capitalized water resource development projects.

The following broad conclusions and recommendations have been drawn from the study:

i. The conceptual principle based on the re-interpretation of the physical meanings of the parameters is sound and the relationships developed are hydrologically meaningful. It is thus possible to develop physically-based procedures to directly quantify parameters of the Pitman model. There is therefore sufficient scope to use some of the physical data obtainable in the region (e.g. maps on geology, soils, topography, etc) in the estimation of the model's parameters. The challenge has been to identify conceptually sensible approaches that can exploit the measurable physical basin attributes. This study has shown the potential of using these attributes in the parameterization of the Pitman model.

ii. The conceptualization, and therefore the parameter estimation procedures developed in this study seem to be quite successful and are at least as good as other regionalization procedures.

iii. This is true even without highly accurate physical basin property data. However, this situation could change with improved availability and access to appropriate physical data in the future through the use of remote sensing and GIS technologies. If these estimation procedures are adopted for use, this will provide an added incentive for the collection and availability of the relevant data.

iv. Using the currently available physical property data involves uncertainties. Improved quantification of these properties would greatly improve the objective application of the estimation procedures.

v. On the strength of the initial results of this first stage in the development of revised physically-based parameter estimation procedures there is need to extend the same approach to the rest of the free parameters of the model. The target parameters are those associated with the interception (i.e. PI1, PI2, and FF) and the evaporation (i.e. PEVAP and R) losses.

Initial intuitive expectations are that these would be related to the basin vegetation cover and rooting depth characteristics.

vi. A significant issue that arises from the appraisal of the "good or better" results from the simulations in the South Africa basins is the sensitivity of the model predictions to variations in parameter values. Should a sensitivity analysis exercise reveal that river flow predictions are insensitive to the revised parameters, then a range of values possibly exists that would result in similar model performance. The implications are that the parameter estimation procedures discussed in this study could not fail as long as the estimated values are within this range. On the other hand, it is also possible that the results are attributable to better estimates of values for a single sensitive parameter. A sensitivity analysis exercise should provide valuable insights into any parameters that need prioritizing for further research or for enhanced data collection to facilitate regionalization. Evidently, there is scope for further work in order to reduce uncertainty in the estimation methods and potential equifinality issues.

vii. With most of the parameters estimated, further work would be required to investigate, identify and quantify the uncertainty associated with the estimation methods and how these impact on the resultant simulations.

viii. When estimation procedures have been developed for the full parameter set of the Pitman model, it will be necessary to establish formal guidelines for their use in the region. These will be expected to provide information on the kind of physical property data required, where they may be accessed from and how to use them. These guidelines should include suggestions for incorporating uncertainty analysis into the parameter estimation approach and therefore into the resulting model simulations.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986. An introduction to the European Hydrological System – Systeme Hydrologique Europeen, SHE. **Journal of Hydrology**, vol. 87, pp. 45–77.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M., 1998. Crop Evapotranspiration: guidelines for computing crop water requirements. Irrigation and drainage paper no. 56. FAO, Rome, Italy.

Ao, T., Ishidaira, H., Takeuchi, K., Kiem, A. S., Yoshitari, J., Fukami, K. and Magome,J., 2006. Relating BTOPMC model parameters to physical features of MPOEX basins. **Journal of Hydrology**, vol. 320, pp. 84–102.

Arnold, J.G. and Allen, P.M., 1996. Estimating hydrologic budgets for three Illinois watersheds. **Journal of Hydrology**, vol. 176, pp. 57–77.

Bailey A.K. and Pitman, W.V., 2005. The Water Resources 2005 Project (WR2005). **Proceedings of the 12th SANCIHAS symposium**, 5 – 7 September 2005, Eskom Convention Centre, Midrand, South Africa.

Baron, J., Seward, P., and Seymour, A. (1998). The groundwater harvest potential map of the Republic South Africa. Technical Report No. GH3917, Directorate of Geohydrology, Department of Water Affairs and Forestry, Pretoria, South Africa.

Basson, M.S., Van Niekerk, P.H. and Van Rooyen, J.A., 1997. Overview of Water Resources Availability and Utilisation in South Africa. Department of Water Affairs and Forestry, Pretoria. 72 pp.

Bathurst, J.C., 1986. Physically based distributed modeling of an upland catchment using the S.H.E.. **Journal of Hydrology**, vol. 87, pp 79–102.

Bergstrom, S. (1991). Principles and confidence in hydrological modelling. **Nordic Hydrology**, vol. 22, pp. 123–136.

Bergstrom, S., 2006. Experience from applications of the HBV hydrological model from the perspective of prediction in ungauged basins. In: Vazken, A., Hall, A. Chahinian, N. and Schaake, J. Large sample basin experiments for hydrological model parameterization: results of the Model Parameter Experiment – MOPEX. **IAHS Publ.** 307, pp. 97-107.

Bergstrom, S. and Forsman, A., (1973). Development of a conceptual deterministic rainfall-runoff model. **Nordic Hydrology**, vol. 4, pp. 147–170.

Beven, K., 1989. Changing ideas in hydrology – the case of physically based models. **Journal of Hydrology**, vol. 105, pp. 157–172.

Beven, K.J., 1993. Prophecy, reality and equifinality in distributed hydrological modelling. **Advances in Water Resources**, vol. 16 (1), pp. 41–51.

Beven, K.J., 2001. **Rainfall-runoff modeling: The primer**. Wiley & Sons Ltd, Chichester, West Sussex England.

Beven, K. J. and Binley, A. M., 1992. The future of distributed models: model calibration and uncertainty prediction. **Hydrological Processes**, vol. 6, pp. 279–298.

Beven, K. J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. **Journal of Hydrology**, vol. 249, pp. 11–29.

Beven, K.J. and Kirkby, M.J., 1979. A physically-based, variable contributing area model of basin hydrology. **Hydrological Sciences Bulletin**, vol. 24(1), pp. 43-69.

Boughton, W. and Chiew, F., 2006. Estimating runoff in ungauged catchments from rainfall, PET and the AWBM model. **Environmental Modelling & Software**, vol. 22(4), pp. 476–487.

Brakensiek, D.L., 1977. Empirical and simplified models of the infiltration process. Report for the USDA-ARS infiltration research planning workshop, St. Louis, MO; United States of America.

Bredenkamp, D.B., Botha, L.J., van Tonder, G.J., and van Rensburg, H.J., 1995. Manual on quantitative estimation of groundwater recharge and aquifer storativity. Water Research Commission Report TT 73/95.

Burke, J.J., Jones, M.J. and Kasimona, V., 1994. Approaches to integrated water resource development and management of the Kafue basin, Zambia. In: C. Kirkby and W.R. White (Eds), **Proceedings of the Integrated River Basin Development International conference**. John Wiley & sons, Chichester.

Burnash, R.J., Ferrul, R.L. and Mcguire, R.A., 1973. A Generalized Stream Flow Simulation System. Conceptual Modeling for Digital Computers. US NWS, Sacramento, CA.

Chiew, F.H.S. and Siriwardena, L., 2005. Estimation of SYMHYD parameter values for application in ungauged catchments. In Zerger, A. and Argent, R.M. (Eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 2883-2889 ISBN: 0-9758400-2-9. http://www.mssanz.org.au/modsim05/papers/chiew_2.pdf

Clarke, R.T., 1973. A review of some mathematical models in hydrology, with observations on their calibration and use. **Journal of Hydrology**, vol. 19, pp. 1–20.

Conrad, J., 2005. Preparation and production of a series of GIS- based maps to identify areas where groundwater contributes to baseflow (WRC Project No. K5/1498). GEOSS Report No. G2005/02-1.

Corwin, D.L. 1996. GIS application of deterministic solute transport models for a regional scale assessment of non-point pollutants in the vadose zone. In: Corwin, D.L. & Loague, K. (Eds). Applications of GIS to the modeling of non-point source pollutants modeling in the vadose zone, SSSA Special Publication No. 48, **Soil Science Society of America**, Madison, Wisconsin.

Cosby, B.J., Hornberger, G.M., Clapp, R.B. & Ginn, T.R. (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. **Water Resources Research**, vol. 20(6), pp. 682-690.

Crawford, N. H. and Linsley, R. K., 1966. "Digital Simulation in Hydrology: Stanford Watershed Model IV," Technical Report, Vol. 39. Department of Civil Engineering, Stanford University, Palo Alto, California, USA.

Dawdy, D.R. and O'Donnell, T., 1965. Mathematical models of catchment behaviour. **Journal of Hydraulic Division**, Am. Soc. Civ. Eng., vol. 91 (HY4), pp 123-137.

De Bruine, B.A., Grobler, A.J. and Crerar, S.E., 1993. Evaluation of a specialized rainfall-runoff model for catchments in Namibia. **Proc. 6th South African National Hydrology Symposium**, Pietermaritzburg, Sept. 1993, pp. 225–233.

DeFries, R., Townshend, J.R.G. and Hansen, M., (1999). Continuous fields of vegetation characteristics at the global scale at 1 km resolution. **Journal of Geophysical Research**, vol. 104, pp. 16911–16925.

De Groen, M.M., 2002. Modelling interception and transpiration at monthly time steps: Introducing daily variability through Markov chains. IHE, Delft, the Netherlands.

De Groen, M.M., and H.H.G. Savenije, 2006. A monthly interception equation based on the statistical characteristics of daily rainfall. **Water Resources Research**, vol. 42 (W12417), pp. 1-10.

Department of Mines (South Africa), 1970. Geological map of the Republic of South Africa and the kingdoms of Lesotho and Swaziland. Pretoria, South Africa.

Department of the Surveyor-General (Zimbabwe Rhodesia), 1979. Zimbabwe Rhodesia 1:1000 000 Soil Map Edition 2. Government Printer, Salisbury, Zimbabwe Rhodesia.

Department of Water Affairs and Forestry (DWAF, South Africa), 2005. Groundwater Resource Assessment II. Pretoria, South Africa.

Direccao Nacional de Geologica (Ministerio dos Recursos Minerais, Mocambique), 1983. Carta Geomorphologica, Republica Popular De Mozambique. Maputo,Mozambique.

Duan, Q., Sorooshian, S. and Gupta, V.K., 1992. Effective and efficient global optimization methods for conceptual rainfall-runoff models. **Water Resources Research**, vol. 28(4), pp. 1015–1031.

Food and Agriculture Organization of the United Nations (FAO), 2003. World Reference Base Map of World Soil Resources, 1:25 000 000. FAO, Rome, Italy.

Fernandez, W., Vogel, R.M. and Sankarasubramanian, A., 2000. Regional calibration of a watershed model. **Hydrological Sciences Journal**, vol. 45(5), pp. 689–707.

Franks, S., 2002. Integrating models, methods and measurements for prediction in ungauged basins. In: Hubert, P., Schertzer, D., Takeuchi, K. and Koide, S., Editors, 2002. Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off meeting held in Brasilia, 20–22 November 2002. **IAHS Publ**. 309, 2007. pp. 13-21.

Gan, T.Y., Dhlamini, E.M., and Biftu, G.B., 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. **Journal of Hydrology**, vol. 192, pp. 81-103.

Görgens, A. H. M., 1983. Conceptual modelling of the rainfall-runoff process in semi- arid catchments. Report No. 1/83, Hydrological Research Unit, Department of Geography, Rhodes University.

Görgens, A. H. M., 1984. Reliability of calibration of a monthly rainfall-runoff model: the semi-arid case. **Hydrological Sciences Journal**, vol. 28(4), pp. 485– 498.

Greco, R., 2002. Preferential flow in macro-porous swelling soil with internal catchment: model development and applications. **Journal of Hydrology**, vol. 269, pp. 150–168.

Hall. R. L., 2003. Interception loss as a function of rainfall and forest types: stochastic modelling for tropical canopies revisited. **Journal of Hydrology**, vol. 280, pp. 1-12.

Hatton, T., 1998. Catchment scale recharge modeling. The basics of recharge and discharge Part 4. CSIRO Publishing.

Hogue, T.S., Sorooshian, S. and Gupta, H., 2000. A multistep automatic calibration scheme for river forecasting models. **Journal of Hydrometeorology**, vol. 1, pp. 524–542.

Hughes, D.A., 1982. Conceptual catchment model parameter transfer studies using monthly data from the southern Cape coastal lakes region. HRU, Report No. 1/82. Department of Geography, Rhodes University, South Africa.

Hughes, D. A., 1993a. Variable time intervals in deterministic hydrological models. **Journal of Hydrology**, vol. 143, pp. 217–232.

Hughes, D. A., 1993b. Soil moisture and runoff simulations using four catchment rainfall-runoff models. **Journal of Hydrology**, vol. 158, pp. 381–404.

Hughes, D. A., 1994. Soil moisture and runoff simulations using four catchment rainfall-runoff models. **Journal of Hydrology**, vol. 158, pp. 381-404.

Hughes, D.A., 1997. Southern African "FRIEND" – The application of rainfall-runoff models in the SADC region. Water Research Commission Report No. 235/1/97.

Hughes, D.A., 2002. The development of an information modelling system for the regional water resource assessments. In: H.J. van Lanen and S Demuth (Eds), FRIEND 2002 –Regional Hydrology: Bridging the gap between research and practice. Proc. 4^{th} FRIEND Conf., Cape Town, March 2002. **IAHS Publ**. 274, pp. 43–49.

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

Hughes, D.A., 2004b. Three decades of hydrological modelling research in South Africa. **South African Journal of Science**, vol. 100, pp. 638–642.

Hughes, D.A. and Forsyth, D., 2006. A generic database and spatial interface for the application of hydrological and water resource models. **Computers and Geosciences**, vol. 32, pp. 1389-1402.

Hughes, D. A. and Metzler, W., 1998. Assessment of three monthly rainfall–runoff models for estimating the water resource yield of semiarid catchments in Namibia. **Hydrological Sciences Journal**, vol. 43 (2), pp. 283–297.

Hughes, D.A. and Parsons, R., 2005. Groundwater Pitman Model Version 3 – Model description. **Proceedings of the 12th South African National Hydrology Symposium**, 5- 7 September 2005, Eskom Convention Centre. Midrand, Gauteng, South Africa.

Hughes, D. A., and Sami, K., 1993. The Bedford catchments. An introduction to their physical and hydrological characteristics. Report to the Water Research Commission. IWR Report No. 1/93, Rhodes University.

Hughes, D. A., and Sami, K., 1994. A semi-distributed, variable time interval model of catchment hydrology – structure and parameter estimation procedures. **Journal of Hydrology**, vol. 155, pp. 265-291.

Hughes, D.A., Murdoch, K.A. and Sami, K., 1994. A hydrological model application system- a tool for integrated river basin management. In: Kirby, C., White, W.R. (Eds), **Integrated River Basin Development**. Wiley, Chichester, UK, pp. 397-406.

Hughes, D.A., Forsyth, D. and Watkins, D.A., 2000. **An integrated software package for the analysis and display of hydrological or water resources time series data.** Water Research Commission Report No. 867/1/2000, Pretoria, South Africa.

Hughes, D.A., Andersson, L., Wilk, J. and Savenije, H.H.G., 2006. Regional calibration of the Pitman model for the Okavango River. **Journal of Hydrology**, vol. 331, pp. 30–42.

Ibbitt, R.P., 1970. Systematic parameter fitting for conceptual models of catchment hydrology. PhD thesis, University of London.

International Water Management Institute (IWMI), 2007. IWMI Online ClimateSummary Service. www.lk.iwmi.org/WAtlas/AtlasQuery.htm.

International Water Management Institute (IWMI), 2007. Integrated Database Information System. http://dw.iwmi.org/dataplatform/Links.aspx.

Inoue, H., 1993. Lateral water flow in a clayey agricultural field with cracks. **Geoderma**, vol. 59, pp. 311–325.

Jakeman, A.J., Littlewood, I. G. and Whitehead, P. G., 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. **Journal of Hydrology**, vol. 117(1-4), pp 275-300.

Klemes, V., 1986. Operational testing of hydrological simulation models. **Hydrological Sciences Journal**, vol.31 (1), pp. 13-24.

Kokkonen, T., 2003. Rainfall-runoff modelling – comparison of modelling strategies with a focus on ungauged predictions and model integration. PhD thesis. Helsinki University of Technology, Finland.

Kokkonen, T.S. and Jakeman, A.J., 2001. A comparison of metric and conceptual approaches in rainfall-runoff modeling and its implications. **Water Resources Research**, vol. 37(9), pp. 2345-2352.

Kostiakov, A.N., 1932. On the dynamics of the coefficient of water percolation in soilsand on the necessity for studying it from a dynamic point of view for purposes of amelioration**. Trans. 6th Int. Soc. Soil Science**, Russian Part A, pp. 17- 21.

Linsley, R. K., 1981. Rainfall-runoff models – an overview. In: Singh V.P. (Ed), Rainfall-runoff relationships. **Water Resources Publications**, pp. 3-22.

Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. **Journal of Hydrology**, vol. 235, pp. 276–288.

Mantovan, P. and Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. **Journal of Hydrology**, vol. 330, pp. 368–381.

Mazvimavi, D. 2003. Estimation of flow characteristics of ungauged catchments. Wageningen University PhD Thesis and International Institute for Geo-Information and Earth Observation, ITC Dissertation Number 107, Enschede, The Netherlands.

Merz, R. and Bloschl, G., 2004. Regionalization of catchment model parameters. **Journal of Hydrology**, vol. 287, pp. 95–123.

Meyboom, P. 1967. Mass transfer studies to determine the groundwater regime of permanent lakes in hummocky moraine of western Canada. **Journal of Hydrology**, vol. 5. pp. 117-142.

Middleton, B.J., Lorentz, S.A., Pitman, W.V. and Midgley, D.C., 1981. Surface water resources of South Africa, Vols. I-VI, Report No. 12/81, Hydrological Research Unit, University of the Witwatersrand, Johannesburg, South Africa.

Midgley, D.C., Pitman, W.V. and Middleton, B.J., 1994. Surface water resources of South Africa 1990, Volumes I to IV, WRC Report No. 298/1.1/94 to 298/6.2/94. Water Research Commission, Pretoria, South Africa.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. **Hydrological Sciences Journal**, vol. 30(2), pp. 273-297.

Mostert, A.C., McKenzie, R.S. and Crerar, S.E., 1993. A rainfall/runoff model for ephemeral rivers in an arid or semi arid environment. **Proc. 6th South African National Hydrology Symposium**, Pietermaritzburg, Sept. 1993, pp. 219–224.

Mwelwa, E.M., 2004. The application of a monthly time step Pitman rainfall-runoff model to the Kafue river basin of Zambia. Rhodes University. Unpublished Thesis.

Nash, J.E. and Sutcliffe, J., 1970. River flow forecasting through conceptual models Part 1. A discussion of principles. **Journal of Hydrology**, vol. 10, pp. 282–290.

Nathan, R.J. and McMahon, T.A., 1990. Identification of homogeneous regions for the purposes of regionalization. **Journal of Hydrology**, vol. 121, pp. 217-238.

Ndiritu, J.G. and Daniell, T.M., 2001. An improved genetic algorithm for rainfall-runoff model calibration and function optimization. **Mathematics & Computer Modelling**, vol. 33, pp. 695-706.

Nicholson, S.E., Kim, J. and Ba, M.B., 1997. The mean surface water balance over Africa and its interannual variability. **Journal of Climate**, vol. 10, pp. 2981-3002.

Nielsen, S.A. and Hansen, E. (1973). Numerical simulation of the rainfall runoff process on a daily basis. **Nordic Hydrology**, vol. 4, pp. 171–190.

Oyebande, L., 2001. Water problems in Africa – how can sciences help? **Hydrological Sciences Journal**, vol. 46(6), pp. 947-961.

Pitman, W.V., 1973. A mathematical model for generating monthly river flows from meteorological data in South Africa. Hydrological Research Unit, Univ. of the Witwatersrand, Report No. 2/73.

Pitman, W.V., 1993. Simulation of run-of-river schemes using monthly data. **Proceedings of the 6th National Hydrological Symposium**, Pietermaritzburg.

Porter, J.W. and McMahon, T.A., 1971. A model for simulation of streamflow data from climatic records. **Journal of Hydrology**, vol. 13, pp. 297–324.

Rawls, W.J., Brakensiek, D.L. & Saxton, K.E. (1982) Estimations of soil water properties. **Transactions of the ASAE**, vol. 25(5), 1316-1320.

Razack, M and Lasm, T, 2006. Geostatistical estimation of the transmissivity in a highly fractured metamorphic and crystalline aquifer (Man-Danane Region, Western Ivory Coast). **Journal of Hydrology**, vol. 325, pp. 164–178.

Refsgaard, J.C. and Storm, B., 1995. MIKE SHE. In: V.P. Singh (ed). Computer models for watershed hydrology. **Water Resources Publications**, pp. 809– 846.

Rhodesia Geological Survey, 1971. Rhodesia geological 1:1000 000, Sixth edition. Rhodesia Geological Survey, Salisbury, Rhodesia.

Rosegrant, M.W. and Perez, N.D., 1997. Water resources development in Africa: a review and synthesis of issues, potentials, and strategies for the future. EPTD Discussion Paper 28, International Food Policy Research Institute, Departmentof Applied Economics, University of Minnesota.

Rutter, A.J., Morton, A.J. and Robins, P.C., 1975. A predictive model of rainfall interception in forests, II. Generation of the model and comparison with observations in some coniferous and hardwood stands. **Journal of Applied Ecology**, vol. 12, pp. 367-380.

Sawunyama, T., Senzanje, J. and Mhizha, A. (2006). Estimation of small reservoir storage capacities in the Limpopo River Basin using geographical information systems (GIS) and remotely sensed surface areas: Case of Mzingwane catchment. **Physics and Chemistry of the Earth**, Parts A/B/C, vol.31(15&16), pp. 935–943.

Sawunyama, T and Hughes, D.A., 2007. Assessment of rainfall and evaporation input data uncertainties on simulated runoff in southern Africa. In: Boegh, E., Kunstmann, H., Wagener, T., Hall, A., Bastidas, L., Franks, S., Gupta, H., Rosbjerg, D. & John Schaake (Eds) Quantification and reduction of predictive uncertainty for sustainable water resources management (Proceedings of Symposium HS2004 at IUGG2007, Perugia, July 2007). **IAHS Publ.** 313, pp. 98-106.

Scanlon, B. R., Healy, R. W. and Cook, P.G., 2002. Choosing appropriate techniques for quantifying groundwater recharge. **Hydrogeology Journal**, vol. 10, pp. 18-39.

Schulze, R.E., 1986. The ACRU model for agrohydrological decision-making: structure, options and application. Proc. Second South African National Hydrology Symp. University of Natal, Pietermaritzburg, Dept. Agric. Engineering, ACRU report 21, pp. 345-362.

Schulze, R.E., 1995. Hydrology and Agrohydrology – A text to accompany the ACRU 3.00 Agrohydrological modelling system. Report to the Water Research Commission by the Department of Agricultural Engineering, University ofNatal. WRC Report No. TT 69/95.

Schulze, R.E. 2006. Soils: Agrohydrological Information Needs, Information Sources and Decision Support. In: Schulze, R.E. (Ed). 2006. South African Atlas of Climatology and Agrohydrology. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/06, Section 4.1.

Schulze, R.E., Hutson, J.L. and Cass, A., 1985. Hydrological characteristics and properties of soils in Southern Africa 2: Soil water retention models. **WaterSA**, vol. 11(3), pp. 129-136.

Schulze, R.E., Maharaj, M, Warburton, M.L., Gers, C.J., Horan, M.J.C and Clark, D.J., 2007. South African Atlas of Climatology and Agrohydrology. University Of Kwazulu-Natal, South Africa.

Siebert, J., 1999. Conceptual runoff models – fiction or representation of reality? PhD thesis. University of Uppsala, Sweden.

SIRI (Soil and Irrigation Research Institute), 1973. Soil map: Republic Of South Africa. Department of Agricultural Technical Services, Pretoria, South Africa.

SIRI (Soil and Irrigation Research Institute), 1987. Land Type Series. Department of Agriculture and Water Supply, Pretoria, RSA, Soil and Irrigation Research Institute. Memoirs on the Agricultural Natural Resources of South Africa.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlebrook, S. and Zehe, E., 2003. IAHS Decade on Prediction in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. **Hydrological Sciences Journal**, vol. 48(6), pp. 857–880.

Smith, M., 1992. Cropwat. A routine for estimation of crop water requirements. FAO Bulletin 57. FAO, Rome, Italy.

Snowy Mountain Engineering Corporation (SMEC), 1987. Southern Okavango integrated Water development Phase 1. Final Report, Technical Study.

Snowy Mountain Engineering Corporation (SMEC), 1991. Botswana National Water Master Plan, Volume 6 – Hydrology. Prepared by SMEC, WLPU and Swedish Geological International AB for the Department of Water Affairs, Botswana.

Sugawara, M., 1995. The development of a hydrological model-tank. Water Resources Publications, LLC; In Kite, G. W. (Ed.), Time and the river: Essays by eminent hydrologists. Highlands Ranch, CO, USA: **Water Resources Publications**, LLC; pp. 201-258.

SWECO, 2004. Development of the Pungwe River Basin Joint Integrated Water Resources Management Strategy. Monograph Report, 2004.

Swedish Meteorological and Hydrological Institute (SMHI), 2000. Streamflow and sediment gauging and modelling in Zimbabwe. GAMZ Main Report.

Todini, E., 1996. The ARNO rainfall-runoff model. **Journal of Hydrology**, vol.175 (1-4), pp. 339-382.

Toth, J., 1963. A theoretical analysis of groundwater flow in small drainage basins. **Journal of Geophysical Research**, vol. 68, pp. 4795-4812.

US Department of Agriculture Soil Conservation Service (1969) Engineering field manual for conservation practices.

Valente, F., David, J.S., and Gash, J.H.C., 1997. Modeling interception loss for two sparse eucalypt and pine forests in central Portugal using reformulated Rutter and Gash analytical models. **Journal of Hydrology**, vol. 190, pp. 141-162.

Wallingford HR, 2003. Handbook for the assessment of catchment water demand and use. DFID, UK.

Wagener, T., Hogue, T., Schaake, J., Duan, Q., Gupta, H., Andreassian, V., Hall, A. and Leavesley, G., 2006. The Model Parameter Estimation Experiment (MOPEX): its structure, connection to other international initiatives and future directions. In: large sample basin experiments for hydrological model parameterization: Results of the Model Parameter Experiment (MOPEX). **IAHS Publ.** 307, pp.339–346.

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V. and Sorooshian, S., 2001. A framework for development and application of hydrological models. **Hydrology & Earth System Sciences**, 5(1), pp.13–26.

Walmsley, R.D., 1991. Stream environments and catchment processes - a SouthAfrican perspective. In: JJ Pigram and BP Hooper (Eds), **Water Allocation to the Environment - Proceedings of an International Seminar and Workshop**. The Centre for Water Policy Research, University of New England, Armidale. pp 45-57.

Warrick, A.W., and Amoozegar-Fard, A., 1979. Infiltration and drain age calculations using spatially scaled hydraulic properties. Water Resources Research, vol. 15, pp. 1116–1120.

Wheater, H.S., 2005. Modeling hydrological processes in arid and semi arid areas – an introduction to the workshop. **Proceedings of the International G-WADI Modelling Workshop**, 28 Feb – 5 Mar 2005, Roorkee, India. www.gwadi.org/shortcourses @ 10:25Hrs, 14/11/2006.

Wheater, H.S., McIntyre, N. and Wagener, T., 2005. Calibration, uncertainty and regional analysis of conceptual rainfall-runoff models. **Proceedings of the International G-WADI Modelling Workshop**, 28 Feb – 5 Mar 2005, Roorkee, India. www.gwadi.org/shortcourses @ 10:25Hrs, 14/11/2006.

World Meteorological Organization (WMO), 1975. Intercomparison of conceptual models used in operational hydrological forecasting. World Meteorological Organization; Operational Hydrology Report No.7. Geneva, Switzerland.

Xu, Y and Beekman, H. E., (Eds), 2003. Groundwater recharge estimation in southern Africa. UNESCO IHP Series No. 64. UNESCO Paris. ISBN 92-9220-000-3.

Zeng, N., Shuttleworth, J.W., and Gash, J.H.C., 2000. Influence of temporal variability of rainfall on interception loss. Part I. Point analysis. **Journal ofHydrology**, vol. 228, pp. 228-241.

Appendices

Appendix 1: Brief descriptions of the physical characteristics of the basins of the study.

Country	Basin code	Gauge	Description
South Africa	B41G	B4H009	Undulating to steep topography, moderate to deep sandy loams; ultra-metamorphics.
	C12D	C2H004	Undulating topography, moderate to deep clayey soils, inter-bedded shales and sandstones.
	G10B	G1H003	Steep topography, shallow, well drained and aerated, porous, sandy loams; unconsolidated sedimentary strata.
	G10A-C	G1H020	Steep rugged topography, shallow, well drained and aerated, porous, sandy loams; unconsolidated sedimentary strata.
	G10E	G1H008	Steep topography, moderately deep, porous sandy loams with some impermeable lenses; unconsolidated sedimentary strata.
	G40J-K	G4H006	Steep topography, shallow, well drained and aerated, porous sandy loams; inter-bedded shales and sandstones.
	G50G	G5H008	Undulating topography, moderate to deep sandy clay loams; unconsolidated sedimentary strata.
	H10A-C	H1H003	Steep, moderately deep sandy loams; Karoo shales and sandstones.
	K20A	K2H002	Steep topography, moderate-deep, permeable, sands; fractured quartzite.
	K40A	K4H003	Steep topography, shallow to moderate loamy sands; fractured granite. Present day impacts of plantations.
	K40C	K4H002	Steep topography, shallow to moderate, sandy loams; fractured quartzite. Present day impacts of pine plantations.
	K60A	K6H001	Steep topography, shallow to moderate sandy loams; fractured quartzites.
	P40A-B	P4H001	Steep topography, moderate to deep clayey loams; fractured granites
	R20C	R2H006	Undulating topography, moderate to deep sandy loams; fractured granites.
	U10A-E	U1H005	Undulating to steep topography, moderate to deep clay loams; Karoo shales, mudstones, sandstones, limestone
	U20B	U2H007	Undulating topography, moderate to deep clays; fractured sedimentary strata.
	U20D	U2H006	Undulating to steep topography, moderate to deep clayey loams; Karoo shales, sandstones, grit and coal.
	V20A	V2H005	Undulating to steep topography, moderate to deep clayey loams; Karoo shales, sandstones, grit and coal.
	V20A-D	V2H002	Undulating to steep topography, moderate to deep clays; fractured granites.
	V31F	V3H009	Undulating topography, moderate to deep clays; porous unconsolidated sedimentary strata.
	V70A	V7H017	Steep topography, shallow to moderate clayey soils, inter-bedded mudstones, shales and sandstones.
	V70B	V7H016	Steep topography, shallow to moderate clayey soils, inter-bedded mudstones, shales and sandstones.
	V70D	V7H012	Steep topography, moderate to deep, clayey soils; inter-bedded mudstones, shales and sandstones.
	W41A-D	W4H004	Undulating to steep topography, moderate to deep sandy loams; Karoo shales, sandstones, grit and coal.
	W52A-C	W5H005	Undulating topography, moderate sandy loam

			soils, inter-bedded sandstones and shales.
	X12A-C	X1H016	Undulating topography, moderate to deep sandy loam soils, gneiss and ultra-metamorphic geology.
	X12J	X1H021	Steep topography, moderate to deep sandy loams; consolidated sedimentary strata.
	X21F-K	X2H015	Undulating to steep topography, moderate to deep sandy clays; porous unconsolidated sedimentary strata.
	X31A	X3H001	Steep topography, moderately deep sandy clay loams; dolomites and limestone.
	X31A-D	X3H006	Steep topography, moderately deep sandy clay loams; dolomites and limestone.
Zimbabwe	BM	B15	Gentle to undulating topography, shallow to moderate shallow loamy sands; fractured gneissic granites.
	BM	B29	Gentle topography, shallow to moderate shallow loamy sands; fractured gneissic granites.
	BT5	B56	Gentle topography, shallow to moderate shallow sandy loams; fractured granites.
	BS3	B77	Gentle topography, moderate to deep sands; Karoo sandstones
	BS3	B78	Gentle topography, moderate to deep sands; Karoo sandstones
	DM7	D27	Gentle topography, shallow to moderate clays; dolerites, mafic meta-volcanics and meta-sediments.
	DM7	D28	Undulating topography, moderate to deep clays; mafic meta-volcanics and meta-sediments
	EO3	E1	Undulating topography, moderate shallow sandy clays; fractured granites.
	EO3	E12	Gentle topography, moderate shallow sands; dolerites, granites- gneissic and massive.
	EM1	E19	Gentle to undulating topography, moderate to deep sands; dolerites, granites- gneissic and massive.
	EO4	E61	Undulating topography, moderate deep sands; dolerites, granites- gneissic and massive.
	ES5	E62	Gentle topography, moderate shallow sands; fractured granites
	EM3	E63	Gentle topography, moderate deep to deep sandy loams; fractured dolerites
	ES3	E114	Gentle topography, moderate shallow sandy loams; granites- gneissic and massive.
	ES3	E115	Undulating topography, moderate deep to deep sandy loams; granites- gneissic and massive.
	EO2	E125	Undulating topography, moderate deep to deep loam clay sands; granites, Umkondo sandstones and quartzites.
	EM2	E136	Gentle topography, moderate deep to deep sandy loams; granites- gneissic and massive.
	EM3	E139	Gentle topography, moderate deep to deep sands; granites- gneissic and massive.
	EM3	E141	Gentle topography, moderate deep to deep sands; granites- gneissic and massive.
	EO4	E147	Gentle topography, moderate shallow sands; fractured granites, meta-volcanics and meta-sediments.
	EM2	E152	Gentle topography, moderate deep to deep silty clay loams; granites- gneissic and massive.
	EO4	E162	Undulating topography, moderate shallow to deep sandy loams; dolerites, granites- gneissic and massive.
	FH	F1	Undulating topography, deep clays; granites- gneissic and massive, dolerites, schists, serpentinites.
	FLS	F7	Undulating topography, deep clays; Umkondo sandstones and quartzites.
	FM2	F10	Undulating topography, deep sandy loams; granites- gneissic and massive.
	FP	F14	Steep topography, deep clays; granites- gneissic and massive
	FB	F18	Undulating topography, moderate shallow to deep loamy

			sands; Umkondo sandstones and quartzites.
	FP	F22	Steep topography, deep clays; granites- gneissic and massive
Mozambique	Unknown	E64	Steep topography, deep sandy clays; granites- gneissic and massive.
	Unknown	E65	Undulating topography, deep sandy clays; fractured granites
	Unknown	E66	Gentle topography, deep silty sandy loams; fractured sedimentary strata.
	Unknown	E72	Undulating topography, moderate to deep sandy loams; fractured granites.
	Unknown	E73	Undulating topography, deep clays; granites- gneissic and massive, dolerites.
Botswana	Unknown	2421	Gentle topography, moderate to deep sands; fractured intrusive igneous granites.
	Unknown	2411	Gentle topography, deep sands, fractured intrusive igneous granites.
Zambia	Unknown	4050	Undulating to steep topography, moderate to deep sandy loams; Kundelungu limestone, shales, banded iron.
	Unknown	4090	Undulating to steep topography, moderate to deep sandy loams; Kundelungu schist, gneiss and granulites.
	Unknown	4120	Undulating topography, moderate to deep sandy loams; fractured granites, gneiss,
	Unknown	4150	Undulating topography, deep sandy loams; fractured granites, limestone, shales, mudstones, slate.
	Unknown	4200	Undulating topography, deep sandy loams; fractured granites, limestone, shales, mudstones, slate.
	Unknown	4340	Undulating topography, deep sandy loams; Kundelungu schist, gneiss and granulites, limestone, shales.

Appendix 2 A summary of the physical property data, estimated physically based parameters and results of model simulations.

South Africa	**B41G**	**C12D**	**G10A_C**	**G10B**	**G10E**	**G40J_K**
MAP (mm)	654	661	1293	1259	649	556
Pot Evapotranspiration (mm/yr)	1500	1580	1515	1515	1635	1430
Basin area (km^2)	448	901	609	46	395	600
WR90 parameters & results						
ST	120	45	270	270	250	250
FT	30	2	100	100	40	4
POW	3.0	3.0	2.0	2.0	2.0	2.0
GW	25	5	20	20	15	15
ZMIN	999	999	0	0	20	20
ZMAX	999	999	400	400	500	350
CE/CE(ln)	0.36/0.56	0.61/0.52	0.85/0.78	0.68/0.59	0.77/0.74	0.32/0.22
R^2/R^2(ln)	0.47/0.57	0.62/0.54	0.89/0.85	0.75/0.88	0.80/0.75	0.48/0.63
%M	-38.0	−16.2	17.9	33.8	−15.1	−5.4
%M(ln)	9.3	142.4	15.3	−183.4	22.7	−68.6
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.14	1.40	1.35	2.10	1.92	1.60
Mean basin slope	0.20	0.10	0.30	0.30	0.25	0.30
Regional GW slope	0.02	0.01	0.05	0.05	0.05	0.05
Drainage vector slope	0.031	0.031	0.042	0.042	0.031	0.031
Mean soil depth (m)	0.65	1.50	1.00	1.00	0.80	0.60
FT soil depth (m)	0.71	1.71	1.13	1.03	1.10	0.67
Soil porosity	0.40	0.35	0.42	0.40	0.40	0.40
Vertical variation factor	0.80	0.80	0.70	0.80	0.63	0.85
Soil Permeability (m/day)	2.438	0.156	0.812	0.270	0.617	1.068
Depth to GW (m)	20	10	12	20	15	15
GW storativity	0.002	0.004	0.002	0.002	0.002	0.003
Unsat transmissivity (m^2/day)	2.0	20.0	5.0	1.4	2.0	4.2
ST_{soil} (mm)	208.0	420.0	294.0	320.0	202.0	204.0
ST_{unsat} (mm)	38.0	0.5	24.0	40.0	30.0	45.0
FT_{soil} (mm/month)	23.83	2.24	22.35	20.10	19.53	20.52
FT_{unsat} (mm/month)	4.24	0.50	17.01	14.11	7.14	12.50
POW	2.0	4.5	2.2	2	1.8	2.1
ZMIN (mm)	20	50	50	20	10	50
ZMEAN (mm)	235	200	225	200	210	250
ZMAX (mm)	450	350	400	380	400	450
CE/CE(ln)	0.42/0.60	0.63/0.61	0.89/0.83	0.74/0.84	0.78/0.76	0.55/0.71
R^2/R^2(ln)	0.44/0.60	0.63/0.64	0.90/0.83	0.75/0.86	0.81/0.77	0.71/0.74
%M	-21.4	0.7	−7.0	−1.8	−16.6	14.3
%M(ln)	-23.2	67.8	−1.2	−42.8	18.7	−29.5

South Africa	G50G	H10A_C	K20A	K40A	K40C	K60A
MAP (mm)	372	590	718	702	926	659
Pot Evapotranspiration (mm/yr)	1430	1650	1400	1400	1400	1540
Basin area (km^2)	382	657	131	72	22	165
WR90 parameters & results						
ST	250	180	100	100	100	100
FT	4	75	50	50	50	25
POW	2.0	2.0	2.0	2.0	2.0	2.0
GW	5	15	50	50	110	18
ZMIN	20	0	0	0	0	30
ZMAX	350	450	200	200	200	600
CE/CE(ln)	0.04/0.27	0.78/0.59	0.66/0.36	0.66/0.57	0.34/0.34	0.61/0.15
R^2/R^2(ln)	0.23/0.47	0.78/0.83	0.79/0.68	0.67/0.68	0.47/0.52	0.64/0.58
%M	29.3	−8.9	42.5	40.7	−31.4	32.2
%M(ln)	−55.8	−38.1	−175.1	−49.8	47.5	−75.4
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.60	1.90	1.20	2.08	3.41	1.91
Mean basin slope	0.08	0.30	0.25	0.30	0.25	0.25
Regional GW slope	0.01	0.03	0.01	0.05	0.04	0.05
Drainage vector slope	1.000	0.031	0.031	0.042	0.031	0.031
Mean soil depth (m)	0.80	1.20	1.20	0.60	0.36	0.29
FT soil depth (m)	0.84	1.53	1.50	0.73	0.43	0.33
Soil porosity	0.42	0.37	0.38	0.39	0.40	0.4
Vertical variation factor	0.44	0.62	0.39	0.80	0.80	0.8
Soil Permeability (m/day)	1.068	0.812	2.438	1.407	2.438	5.564
Depth to GW (m)	8	30	10	30	10	10
GW storativity	0.002	0.002	0.002	0.002	0.002	0.002
Unsat transmissivity (m^2/day)	2.0	2.0	60.0	5.0	3.4	3
ST_{soil} (mm)	148.0	275.0	178.0	187.0	115.0	93
ST_{unsat} (mm)	0.0	6.5	18.0	60.0	20.0	20
FT_{soil} (mm/month)	6.93	42.56	65.83	38.62	53.45	52.38
FT_{unsat} (mm/month)	0.00	6.50	1.30	26.21	21.56	10.66
POW	3.2	1.8	2.2	2.0	2.0	4.5
ZMIN (mm)	50	10	20	10	0	200
ZMEAN (mm)	100	110	160	220	50	500
ZMAX (mm)	550	210	200	250	100	800
CE/CE(ln)	0.36/0.53	0.75/0.72	0.75/0.63	0.66/0.67	0.51/0.55	0.64/0.50
R^2/R^2(ln)	0.63/0.55	0.75/0.82	0.76/0.64	0.63/0.70	0.55/0.57	0.65/0.56
%M	−20.9	−4.6	−11.0	−3.7	0.6	−1.7
%M(ln)	−18.1	−24.7	−18.9	−18.8	−16.5	−0.2

South Africa	P40A_B	R20C	U10A_E	U20B	U20D	V20A
MAP (mm)	599	809	1071	984	1027	1028
Pot Evapotranspiration (mm/yr)	1500	1450	1300	1300	1300	1300
Basin area (km^2)	576	121	1744	358	339	260
WR90 parameters & results						
ST	200	200	100	200		
FT	2	12	50	30	200	100
POW	3.0	3.0	3.0	3.0	30	50
GW	15	6.7	20	15	3.0	3.0
ZMIN	20	45	999	999	22	40
ZMAX	600	600	999	999	999	999
CE/CE(ln)	0.66/0.56	0.27/0.46	0.66/0.76	0.18/0.57	999	999
R^2/R^2(ln)	0.73/0.40	0.64/0.51	0.74/0.85	0.56/0.65	0.36/0.67	0.80/0.49
%M	−38.1	−3.2	−22.4	11.1	0.43/0.73	0.86/0.82
%M(ln)	866.6	26.8	−6.3	6.3	−35.8	−18.5
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.50	1.20	1.20	1.36	1.73	2.57
Mean basin slope	0.18	0.15	0.18	0.15	0.20	0.30
Regional GW slope	0.02	0.01	0.02	0.02	0.02	0.05
Drainage vector slope	0.042	0.042	0.042	0.042	0.031	0.042
Mean soil depth (m)	1.20	1.20	1.20	1.20	1.20	1.20
FT soil depth (m)	1.38	1.38	1.24	1.27	1.38	1.47
Soil porosity	0.38	0.40	0.38	0.39	0.43	0.39
Vertical variation factor	0.80	0.80	0.80	0.80	0.80	0.70
Soil Permeability (m/day)	0.156	0.617	2.438	0.270	1.852	0.156
Depth to GW (m)	20	20	20	30	20	15
GW storativity	0.002	0.002	0.002	0.002	0.002	0.003
Unsat transmissivity (m^2/day)	1.6	1.8	5.0	2.5	3.9	2.1
ST_{soil} (mm)	365.0	384.0	365.0	374.0	413.0	328.0
ST_{unsat} (mm)	35.0	31.0	35.0	50.0	38.0	45.0
FT_{soil} (mm/month)	3.47	9.15	39.32	4.19	52.87	10.58
FT_{unsat} (mm/month)	6.05	5.44	15.12	8.57	12.55	13.79
POW	2.5	2.5	3.0	2.4	3.0	2.5
ZMIN (mm)	50	20	50	10	20	10
ZMEAN (mm)	200	330	175	300	200	170
ZMAX (mm)	350	640	300	590	340	330
CE/CE(ln)	0.70/0.56	0.54/0.52	0.70/0.77	0.52/0.67	0.50/0.74	0.83/0.83
R^2/R^2(ln)	0.71/0.62	0.62/0.58	0.74/0.84	0.54/0.70	0.50/0.75	0.84/0.85
%M	−0.3	2.3	−3.2	−2.8	−11.0	−8.6
%M(ln)	−5.1	−3.1	8.3	13.5	7.1	−6.0

South Africa	V20A_D	V31F	V70A	V70B	V70D	W41A_D
MAP (mm)	956	922	1177	1093	814	943
Pot Evapotranspiration (mm/yr)	1300	1450	1300	1300	1350	1400
Basin area (km^2)	937	148	276	121	196	948
WR90 parameters & results						
ST						
FT	100	120	100	100	120	100
POW	50	15	50	50	30	25
GW	3.0	3.0	3.0	3.0	3.0	2.0
ZMIN	40	2.2	25	20	15	20
ZMAX	999	999	999	999	999	999
CE/CE(ln)	999	999	999	999	999	999
R^2/R^2(ln)	0.76/0.83	0.37/0.61	0.59/0.77	0.69/0.79	0.51/0.62	0.43/0.57
%M	0.77/0.84	0.52/0.64	0.60/0.78	0.70/0.79	0.62/0.73	0.59/0.67
%M(ln)	−7.0	−8.3	−10.0	2.8	11.0	−3.9
Revised physically based parameters & results						
Drainage Density (km/km^2)	2.57	1.62	2.40	1.50	2.34	1.09
Mean basin slope	0.30	0.12	0.30	0.30	0.30	0.20
Regional GW slope	0.05	0.01	0.05	0.01	0.05	0.04
Drainage vector slope	0.042	0.042	0.031	0.020	0.042	0.042
Mean soil depth (m)	1.20	0.90	0.80	0.80	0.80	1.00
FT soil depth (m)	1.47	0.95	0.93	0.97	0.93	1.08
Soil porosity	0.39	0.35	0.32	0.35	0.32	0.35
Vertical variation factor	0.70	0.70	0.80	0.80	0.80	0.60
Soil Permeability (m/day)	0.156	1.407	1.407	0.617	0.356	1.407
Depth to GW (m)	15	15	15	20	15	20
GW storativity	0.003	0.001	0.003	0.002	0.003	0.002
Unsat transmissivity (m^2/day)	2.1	1.0	5.0	1.4	1.0	3.0
ST_{soil} (mm)	328.0	221.0	205.0	224.0	205.0	210.0
ST_{unsat} (mm)	45.0	11.0	45.0	40.0	45.0	40.0
FT_{soil} (mm/month)	10.58	15.59	56.72	16.01	13.98	19.78
FT_{unsat} (mm/month)	13.79	4.08	4.46	2.01	5.90	8.24
POW	2.5	3.5	4.0	2.0	4.0	3.0
ZMIN (mm)	10	20	40	10	30	20
ZMEAN (mm)	170	320	300	225	300	300
ZMAX (mm)	330	620	430	440	550	580
CE/CE(ln)	0.76/0.83	0.48/0.61	0.62/0.80	0.66/0.77	0.60/0.75	0.60/0.73
R^2/R^2(ln)	0.76/0.84	0.52/0.62	0.64/0.81	0.67/0.79	0.62/0.75	0.61/0.77
%M	−3.0	−12.4	−13.4	0.4	5.2	−7.3
%M(ln)	3.7	−19.0	1.4	12.3	8.1	7.0

South Africa	W52A_C	X12A_C	X12J	X21F_K	X31A	X31A_D
MAP (mm)	837	829	1156	967	1243	1182
Pot Evapotranspiration (mm/yr)	1400	1400	1400	1400	1400	1400
Basin area (km^2)	804	581	295	1554	174	766
WR90 parameters & results						
ST	180	150	500	140	600	600
FT	15	24	15	20	60	60
POW	3.0	2.0	2.0	2.0	2.0	2.0
GW	15	18	49	15	60	60
ZMIN	0	999	0	99	0	0
ZMAX	900	999	900	999	800	800
CE/CE(ln)	0.62/0.72	0.24/0.61	0.49/0.25	0.56/0.71	0.67/0.47	0.79/0.74
R^2/R^2(ln)	0.64/0.75	0.48/0.64	0.63/0.72	0.65/0.76	0.74/0.70	0.82/0.82
%M	–12.2	17.3	40.3	14.1	23.2	13.6
%M(ln)	8.2	13.4	46.0	7.5	24.1	7.0
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.60	1.80	1.65	1.18	1.47	1.67
Mean basin slope	0.10	0.12	0.20	0.15	0.25	0.20
Regional GW slope	0.01	0.01	0.01	0.02	0.05	0.03
Drainage vector slope	0.042	0.02	0.042	0.031	0.031	0.042
Mean soil depth (m)	1.00	1.80	1.50	1.60	1.00	1.60
FT soil depth (m)	1.03	1.95	1.56	1.72	1.09	1.69
Soil porosity	0.35	0.38	0.40	0.33	0.38	0.39
Vertical variation factor	0.60	0.55	0.90	0.70	0.72	0.70
Soil Permeability (m/day)	0.812	0.812	0.118	0.617	0.468	1.407
Depth to GW (m)	10	20	18	20	10	25
GW storativity	0.003	0.002	0.004	0.002	0.002	0.002
Unsat transmissivity (m^2/day)	2.5	8.0	1.0	2.0	2.5	4.1
ST_{soil} (mm)	210.0	376.0	540.0	370.0	274.0	437.0
ST_{unsat} (mm)	20.0	3.0	60.0	37.0	20.0	47.0
FT_{soil} (mm/month)	8.05	20.51	3.65	11.28	11.21	47.52
FT_{unsat} (mm/month)	0.40	3.00	4.16	4.39	6.84	17.09
POW	2.5	2.8	3.0	2.5	2.0	2.8
ZMIN (mm)	0	50	50	50	0	0
ZMEAN (mm)	325	350	300	300	375	400
ZMAX (mm)	650	650	550	550	750	800
CE/CE(ln)	0.41/0.66	0.52/0.67	0.55/0.67	0.65/0.79	0.74/0.71	0.82/0.82
R^2/R^2(ln)	0.48/0.68	0.52/0.70	0.62/0.73	0.65/0.79	0.74/0.71	0.83/0.83
%M	–3.0	–2.4	13.5	–4.2	–2.3	1.9
%M(ln)	12.6	6.1	13.2	1.5	3.3	0.5

Zimbabwe	**B15**	**B29**	**B56**	**B77**	**B78**	**D27**
MAP (mm)	535	570	564	610	575	858
Pot Evapotranspiration (mm/yr)	1535	1563	1563	1580	1580	1405
Basin area (km^2)	267	363	645	539	49.2	70
Revised physically based parameters & results						
Drainage Density (km/km^2)	0.96	1.20	1.40	0.30	0.80	1.50
Mean basin slope	0.08	0.07	0.1	0.07	0.1	0.08
Regional GW slope	0.01	0.01	0.01	0.01	0.02	0.01
Drainage vector slope	0.042	0.042	0.02	0.042	0.042	0.042
Mean soil depth (m)	0.61	0.85	0.85	0.60	0.90	1.20
FT soil depth (m)	0.63	0.89	0.89	0.63	0.93	1.27
Soil porosity	0.40	0.40	0.40	0.42	0.41	0.41
Vertical variation factor	0.93	0.80	0.80	1.00	0.80	0.95
Soil Permeability (m/day)	1.407	1.852	3.210	5.564	0.812	0.118
Depth to GW (m)	25	34	20	35	30	12
GW storativity	0.003	0.002	0.002	0.004	0.004	0.030
Unsat transmissivity (m^2/day)	3.2	4.3	4.8	2.8	0.9	1.9
ST_{soil} (mm)	227.0	255.0	272.0	252.0	295.0	467.0
ST_{unsat} (mm)	41.0	28.0	36.0	65.0	87.0	196.0
FT_{soil} (mm/month)	4.11	8.30	23.97	4.44	3.64	1.09
FT_{unsat} (mm/month)	7.74	13.00	8.06	2.12	1.81	6.99
POW	2.8	3.9	3.0	3.5	3.5	2.0
ZMIN (mm)	50	50	50	50	20	20
ZMEAN (mm)	250	250	300	500	340	510
ZMAX (mm)	350	450	550	950	660	1000
CE/CE(ln)	0.73/0.61	0.70/0.56	0.73/0.50	0.50/0.60	0.79/0.74	0.79/0.52
R^2/R^2(ln)	0.74/0.62	0.70/0.57	0.74/0.52	0.64/0.60	0.79/0.76	0.78/0.52
%M	−8.3	−9.9	−0.4	−2.2	−0.5	7.9
%M(ln)	−9.0	−8.3	1.9	−5.2	−10.2	−8.1

Zimbabwe	D28	E1	E12	E19	E61	E62
MAP (mm)	870	1080	1028	1096	841	766
Pot Evapotranspiration (mm/yr)	1405	1328	1404	1520	1374	1419
Basin area (km^2)	223	249	3230	3320	2450	1990
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.50	3.10	1.60	1.10	2.10	0.75
Mean basin slope	0.08	0.12	0.10	0.05	0.12	0.07
Regional GW slope	0.01	0.02	0.01	0.01	0.01	0.01
Drainage vector slope	0.042	0.042	0.042	0.042	0.042	0.042
Mean soil depth (m)	1.20	2.70	1.65	2.10	1.20	0.80
FT soil depth (m)	1.27	2.88	1.72	2.14	1.38	0.84
Soil porosity	0.41	0.40	0.42	0.42	0.41	0.42
Vertical variation factor	0.95	1.00	0.80	1.00	0.80	0.85
Soil Permeability (m/day)	0.118	0.090	0.156	0.205	1.852	1.852
Depth to GW (m)	12	10	25	20	15	10
GW storativity	0.030	0.040	0.050	0.080	0.015	0.006
Unsat transmissivity (m^2/day)	1.9	0.5	0.6	0.9	2.2	3.7
ST_{soil} (mm)	467.0	1080.0	554.0	882.0	394.0	286.0
ST_{unsat} (mm)	196.0	312.0	807.0	320.0	160.0	28.0
FT_{soil} (mm/month)	1.09	5.77	2.58	1.45	38.50	4.93
FT_{unsat} (mm/month)	6.99	3.75	2.50	2.49	11.64	6.99
POW	2.0	3.0	3.2	2.0	2.0	3.5
ZMIN (mm)	20	10	10	50	10	50
ZMEAN (mm)	510	605	605	650	405	400
ZMAX (mm)	1000	1200	1200	1250	800	750
CE/CE(ln)	0.71/0.53	0.63/0.61	0.80/0.80	0.66/0.52	0.68/0.70	0.63/0.70
R^2/R^2(ln)	0.71/0.67	0.70/0.67	0.81/0.82	0.80/0.57	0.71/0.70	0.68/0.71
%M	0.8	–5.2	–5.6	30.0	–2.7	–1.9
%M(ln)	–6.7	14.8	4.8	24.0	2.2	7.8

Zimbabwe	**E63**	**E114**	**E115**	**E125**	**E136**	**E139**
MAP (mm)	795	1037	1037	938	808	787
Pot Evapotranspiration (mm/yr)	1450	1430	1430	1295	1455	1470
Basin area (km^2)	989	197	223	433	635	329
Revised physically based parameters & results						
Drainage Density (km/km^2)	2.28	1.50	2.10	3.12	1.41	1.30
Mean basin slope	0.05	0.08	0.10	0.15	0.06	0.08
Regional GW slope	0.02	0.01	0.02	0.01	0.01	0.01
Drainage vector slope	0.042	0.02	0.042	0.042	0.042	0.042
Mean soil depth (m)	1.20	2.00	1.20	2.00	0.90	1.25
FT soil depth (m)	1.27	2.11	1.27	2.15	0.93	1.28
Soil porosity	0.40	0.42	0.40	0.33	0.40	0.42
Vertical variation factor	0.55	0.80	0.85	0.80	0.85	0.80
Soil Permeability (m/day)	4.226	0.617	0.356	0.356	1.852	3.210
Depth to GW (m)	10	17	12	25	12	10
GW storativity	0.005	0.050	0.005	0.015	0.003	0.004
Unsat transmissivity (m^2/day)	1.0	3.2	0.9	1.8	2.0	1.5
ST_{soil} (mm)	264.0	756.0	408.0	528.0	306.0	420.0
ST_{unsat} (mm)	13.0	729.0	44.0	290.0	13.0	22.0
FT_{soil} (mm/month)	36.62	9.37	5.71	21.47	8.77	46.29
FT_{unsat} (mm/month)	5.75	5.76	5.29	14.15	7.11	4.91
POW	4.0	5.0	2.0	5.0	2.5	3.5
ZMIN (mm)	50	50	20	10	50	10
ZMEAN (mm)	455	625	410	730	300	405
ZMAX (mm)	860	1200	800	1450	550	800
CE/CE(ln)	0.68/0.68	0.46/0.64	0.70/0.61	0.64/0.73	0.71/0.41	0.75/0.22
R^2/R^2(ln)	0.69/0.68	0.54/0.69	0.75/0.61	0.73/0.73	0.71/0.43	0.75/0.53
%M	−9.0	−8.9	−6.0	1.0	−7.4	−0.5
%M(ln)	9.5	−62.6	5.0	10.0	3.2	−8.7

Zimbabwe	**E141**	**E147**	**E152**	**E162**	**F1**	**F7**
MAP (mm)	741	752	844	942	1637	1030
Pot Evapotranspiration (mm/yr)	1520	1381	1470	1325	1314	1320
Basin area (km^2)	2820	318	146	2165	6.5	127
Revised physically based parameters & results						
Drainage Density (km/km^2)	1.78	1.20	1.35	2.00	2.54	3.15
Mean basin slope	0.04	0.05	0.06	0.06	0.20	0.12
Regional GW slope	0.01	0.01	0.01	0.01	0.05	0.05
Drainage vector slope	0.042	0.042	0.042	0.042	0.042	0.042
Mean soil depth (m)	1.30	0.91	1.30	1.60	2.20	2.30
FT soil depth (m)	1.38	0.94	1.38	1.72	2.50	2.90
Soil porosity	0.42	0.42	0.33	0.40	0.32	0.37
Vertical variation factor	0.80	0.95	0.80	0.90	0.85	0.72
Soil Permeability (m/day)	5.564	1.068	0.812	0.812	0.812	0.468
Depth to GW (m)	25	22	10	25	10	7
GW storativity	0.002	0.020	0.004	0.020	0.008	0.050
Unsat transmissivity (m^2/day)	2.0	0.4	2.9	1.2	2.0	20.0
ST_{soil} (mm)	437.0	363.0	343.0	576.0	598.0	613.0
ST_{unsat} (mm)	0.0	88.0	14.0	180.0	80.0	1.6
FT_{soil} (mm/month)	32.75	3.61	5.43	10.06	61.85	30.80
FT_{unsat} (mm/month)	0.00	1.36	9.87	6.05	12.80	1.60
POW	3.7	3.0	3.0	2.5	2.2	3.5
ZMIN (mm)	50	10	0	10	20	10
ZMEAN (mm)	325	445	250	605	285	705
ZMAX (mm)	600	880	500	1200	550	1400
CE/CE(ln)	0.74/0.51	0.86/0.52	0.64/0.33	0.64/0.61	0.86/0.88	0.67/0.65
R^2/R^2(ln)	0.75/0.53	0.86/0.57	0.65/0.36	0.66/0.62	0.86/0.89	0.69/0.66
%M	–9.9	–0.5	–8.4	–8.5	–7.5	–8.2
%M(ln)	11.3	–5.3	–5.9	4.7	3.9	6.7

Zimbabwe & Mozambique	**F10**	**F14**	**F18**	**F22**	**E64**	**E65**
MAP (mm)	1620	1576	970	1835	1472	1092
Pot Evapotranspiration (mm/yr)	1445	1416	1427	1416	1486	1493
Basin area (km^2)	31	85.5	148	641	687	1313
Revised physically based parameters & results						
Drainage Density (km/km^2)	3.00	4.00	2.40	4.00	3.90	1.50
Mean basin slope	0.17	0.25	0.15	0.25	0.22	0.15
Regional GW slope	0.02	0.05	0.02	0.05	0.05	0.03
Drainage vector slope	0.042	0.042	0.042	0.042	0.042	0.042
Mean soil depth (m)	2.60	1.50	1.90	1.80	0.97	1.37
FT soil depth (m)	2.64	1.83	2.11	1.88	1.08	1.69
Soil porosity	0.40	0.40	0.40	0.39	0.40	0.36
Vertical variation factor	1.00	0.90	0.90	0.80	0.80	0.80
Soil Permeability (m/day)	0.356	1.068	0.468	1.068	0.156	2.438
Depth to GW (m)	14	20	17	8	30	25
GW storativity	0.050	0.003	0.010	0.005	0.050	0.002
Unsat transmissivity (m^2/day)	3.0	3.3	3.1	3.0	1.0	2.5
ST_{soil} (mm)	1040.0	540.0	684.0	562.0	310.0	395.0
ST_{unsat} (mm)	598.0	60.0	141.0	40.0	1500.0	45.0
FT_{soil} (mm/month)	28.76	117.54	21.39	120.21	13.37	55.55
FT_{unsat} (mm/month)	22.68	32.76	18.99	29.77	9.83	9.45
POW	3.0	1.0	3.0	2.5	2.3	2.5
ZMIN (mm)	30	20	20	20	50	50
ZMEAN (mm)	565	160	660	160	650	350
ZMAX (mm)	1100	300	1300	300	1250	650
CE/CE(ln)	0.63/0.87	0.52/0.62	0.60/0.68	0.66/0.52	0.71/0.81	0.79/0.91
R^2/R^2(ln)	0.73/0.88	0.52/0.67	0.64/0.68	0.66/0.52	0.79/0.83	0.81/0.92
%M	−2.8	−6.6	−10.6	−0.1	8.9	−9.7
%M(ln)	5.0	−3.1	−8.6	2.6	1.6	−0.1

Mozambique & Botswana	**E66**	**E72**	**E73**	**2421**	**2411/2511**
MAP (mm)	968	1168	1646	459	468
Pot Evapotranspiration (mm/yr)	1380	1526	1479	1555	1555
Basin area (km^2)	15046	2700	1100	1320	2250
Revised physically based parameters & results					
Drainage Density (km/km^2)	1.50	0.90	1.50	0.10	0.09
Mean basin slope	0.08	0.12	0.15	0.05	0.05
Regional GW slope	0.01	0.01	0.02	0.01	0.01
Drainage vector slope	0.031	0.042	0.042	0.042	0.042
Mean soil depth (m)	1.95	2.30	2.00	1.20	3.00
FT soil depth (m)	1.99	2.38	2.13	1.38	3.26
Soil porosity	0.40	0.39	0.45	0.41	0.41
Vertical variation factor	0.85	0.95	1.00	0.80	0.80
Soil Permeability (m/day)	5.564	0.468	0.270	1.407	0.468
Depth to GW (m)	20	25	20	40	55
GW storativity	0.010	0.020	0.057	0.003	0.020
Unsat transmissivity (m^2/day)	1.0	1.7	1.5	2.5	0.6
ST_{soil} (mm)	663.0	852.0	900.0	394.0	984.0
ST_{unsat} (mm)	140.0	355.0	949.0	24.0	220.0
FT_{soil} (mm/month)	79.67	7.21	7.75	0.58	0.41
FT_{unsat} (mm/month)	2.79	3.86	5.67	0.62	0.14
POW	5.0	3.0	2.6	3.0	3.0
ZMIN (mm)	50	50	0	20	50
ZMEAN (mm)	650	650	350	500	650
ZMAX (mm)	1250	1250	700	980	1250
CE/CE(ln)	0.37/0.70	0.50/0.82	0.66/0.83	0.61/0.22	0.50/0.30
R^2/R^2(ln)	0.37/0.72	0.55/0.82	0.75/0.83	0.61/0.27	0.58/ 0.32
%M	−0.6	−2.5	5.5	−9.2	5.0
%M(ln)	2.9	2.5	1.0	−46.3	4.7

Zambia	**4050**	**4090**	**4120**	**4150**	**4200**	**4340**
MAP (mm)	1283	1236	1236	1236	1214	1200
Pot Evapotranspiration (mm/yr)	1640	1504	1604	1504	1464	1464
Basin area (km^2)	4999	7148	869	9195	11655	8708
Revised physically based parameters & results						
Drainage Density (km/km^2)	2.40	2.10	2.40	2.10	1.50	1.50
Mean basin slope	0.20	0.20	0.15	0.15	0.15	0.15
Regional GW slope	0.04	0.04	0.02	0.03	0.03	0.02
Drainage vector slope	0.031	0.031	0.042	0.042	0.042	0.042
Mean soil depth (m)	1.60	1.80	1.60	1.80	2.00	1.80
FT soil depth (m)	1.67	1.88	1.67	1.95	2.08	2.00
Soil porosity	0.42	0.42	0.40	0.40	0.40	0.40
Vertical variation factor	0.90	0.90	0.90	0.80	0.80	0.80
Soil Permeability (m/day)	1.407	1.852	2.438	0.812	1.407	0.468
Depth to GW (m)	25	25	25	28	20	25
GW storativity	0.020	0.020	0.030	0.030	0.040	0.050
Unsat transmissivity (m^2/day)	2.8	2.5	2.0	3.7	2.5	3.2
ST_{soil} (mm)	605.0	680.0	576.0	576.0	640.0	576.0
ST_{unsat} (mm)	500.0	500.0	624.0	757.0	721.0	1040.0
FT_{soil} (mm/month)	67.86	87.51	88.22	29.91	39.41	12.64
FT_{unsat} (mm/month)	12.50	9.77	12.10	19.58	9.45	12.10
POW	2.0	3.3	4.5	4.5	4.2	5.0
ZMIN (mm)	200	200	100	100	100	200
ZMEAN (mm)	550	700	500	650	600	800
ZMAX (mm)	900	1200	1200	1200	1200	1400
CE/CE(ln)	0.60/0.76	0.75/0.86	0.67/0.77	0.78/0.84	0.75/0.84	0.73/0.75
R^2/R^2(ln)	0.64/0.83	0.75/0.86	0.68/0.76	0.79/0.84	0.79/0.86	0.73/0.76
%M	9.0	−0.3	0.9	2.1	−1.3	0.3
%M(ln)	9.3	2.0	1.5	0.6	2.5	2.8