Satellite data as complementary information for hydrological modelling

H. C. WINSEMIUS

Satellite data as complementary information for hydrological modelling

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J. T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op woensdag 30 september 2009 om 12.30 uur

door

Hessel Constantijn WINSEMIUS

Civiel ingenieur
geboren op 21 januari 1980 te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. H. H. G. Savenije

Samenstelling promotiecommissie:

Rector Magnificus, *voorzitter*

Prof. dr. ir. **H. H. G. Savenije**, *Technische Universiteit Delft, promotor*

Prof. Dr.-Ing. habil. **R. Klees**, *Technische Universiteit Delft*

Prof. dr. **B. J. J. M. van den Hurk**, *Universiteit Utrecht*

Prof. dr. ir. **P. van der Zaag**, *UNESCO-IHE – Technische Universiteit Delft*

Prof. dr. **P. A. Troch**, *University of Arizona*

Prof. dr. **A. Montanari**, *Università di Bologna*

Prof. dr. **G. Blöschl**, *Technische Universität Wien*

Prof. dr. **W. G. M. Bastiaanssen**, *Technische Universiteit Delft*

First edition: 2009

Title: **Satellite data as complementary information for hydrological modelling**

Preface

I have always found rainfall intriguing. I remember from my childhood that I sheltered under a chestnut tree in the garden of my parents' second house on the island of Ameland, to witness a heavy thunderstorm (not realising this might actually be a bit dangerous). It is really astonishing to see how dependent rainfall occurrence is on the interplay of sometimes highly localized non-linearities and feedbacks. During my field work in Africa, I witnessed a ferocious storm originating a few kilometres away, while the sun was shining on my head. Apparently, the conditions for uplift of moist air were just not met at my position, causing spatially extremely variable conditions. The unpredictability and variability of these conditions and the feedbacks that generate them makes rainfall difficult to predict. One may predict the conditions, but often not the event itself.

Going from rainfall to hydrology, similar non-linearities occur in rainfall-runoff processes. In some catchments, a large rainfall event may result in zero discharge if conditions are dry. It will require a number of events for a small amount of the rainfall to result in discharge. It is important to realise that such non-linearities are greater in areas with high evaporation potential and low rainfall amounts, where only a small proportion of the rainfall eventually ends up in the river. The cascading of non-linearities in both the atmosphere and runoff generating processes is one of the main reasons why these semi-arid areas are more frequently hit by floods and droughts and why a higher sensitivity to climate change may be expected. My good friend Michael Mutale, who happens to live in a semi-arid region in Zambia, has told me many stories about both failure of harvests due to persistent dry periods and death tolls due to floods in the Zambezi river. But at the same time water is the source of health and well being. All over the world, livelihoods, economies, cultures and nature depend on water as primary resource. The dependence on water on the one hand, and, threats created by water on the other, have almost everywhere impact on human life and nature. And yet, the quantification of water resources and the prediction of extreme events remains difficult. In fact, in large parts of the world, little is understood about the complexity and the quantification of the mechanisms that cause runoff, which in turn often is caused by a lack of data.

If we do not understand the present behaviour, it is even more difficult to predict the effects of future change. The behaviour of a catchment is a complex interplay at different spatial and temporal scales between climate, geology, soil characteristics and ecosystems. It is therefore inavoidable that if climate changes, the ecosystem will adapt and previously dominating species may be suppressed while new species will become dominant. Because of this, soil texture and root channel distribution will change and at longer time scales, erosion processes may reshape the catchment. It will change the topology and structure of the substrate and as a result the landscape as a whole. In order to predict such feedbacks, a proper knowledge of currently governing hydrological processes should be the first step.

The lack of predictive power in ungauged basins is one of the main reasons why the International Association of Hydrological Sciences (IAHS) initiated the decade of 'Predictions in Ungauged Basins' (PUB, Sivapalan, 2003). In the past years, a considerable amount of studies has been performed in the framework of PUB that concentrate on the development of models and prediction methods in ungauged basins. More specifically, a lot of work has focussed on model regionalisation, parameter transferability, the use of soft information, uncertainty estimation methods with scarce data and the use of new observations techniques. Ironically, most published studies that I noticed so far on this subject deal with the development of these methods in basins where, first of all, a lot of measurements are available, and second, different climate conditions govern the hydrology than in the regions where real ungauged conditions occur. To demonstrate a new method, observations are removed and model development is performed without the used data. It seems to me that this leads to the self-fulfilling prophecy that the case study is always too perfect to be true and suiting unrealistic conditions. As a result, modellers tend to overlook that they make certain assumptions which may not be applicable in a real ungauged case.

This is the reason why I have concentrated my efforts on a real ungauged basin, rather than an artificial ungauged basin. Hence, I was forced not to 'cheat' and to explore what can be done if all data problems occur simultaneously instead of one at a time. Furthermore, the spatial scale of the selected study areas chosen, is consistent with the spatial scales of real-world problems. It took a while before I dared to release the paradigm of classical model calibration, but once I did this I started realising that any data has potential value. It is up to the modeller to convert this value to useful information. It turns out that satellite data (most of them free of charge) can be used to find realistic approximations of catchment configuration, especially when these can be combined with a few ground measurements of rainfall and discharge. Hence a modeller can increase his understanding of catchment functioning. I hope my work helps to answer real-world demands of hydrologists and that this PhD thesis offers a valuable contribution to the IAHS decade for Prediction in Ungauged Basins.

Summary

Many catchments in the world are scarcely gauged. The little measurements that are available in these regions are often uncertain, scarce, intermittent or non-concomitant. This jeopardises the construction of hydrological models, that in the short term are needed to predict droughts, floods and availability of water, and in the long term the effects of changes in the land cover or of the climate. Fortunately, an increasing amount of data, having a possible use in the development, calibration and validation of hydrological models, is made available to the public. In particular satellite data are very interesting. Important variables such as rainfall, evaporation, radiation, soil moisture and water storage can nowadays be estimated on the basis of raw satellite observations.

This thesis describes the development and application of methods that allow to combine the scarce data, available in poorly gauged catchments, with expert knowledge and modern satellite data, with the purpose to conceptualise, calibrate and validate hydrological models. The satellite information is here considered to be 'orthogonal' information, as it provides an independent view on the functioning of a catchment. The methods have been developed in such a way, that the information stored in all the observations is optimally used, while considering the uncertainties in both ground measurements and satellite data in an objective way. It is demonstrated that in this way, even poor quality data may be used as model input and output. The complementary satellite data used, are comprised of satellite based rainfall, gravity measurements from the satellite mission Gravity Recovery And Climate Experiment (GRACE) and satellite based evaporation estimates. The methods have been developed based on a number of case studies in the catchment of the Zambezi. This is a scarsely gauged catchment, offering an excellent opportunity to directly bring the methods developed to practice.

Chapter 4 provides an analysis of the errors in the GRACE observations and how these possibly may influence the hydrological storage signal and thus its application in the development of hydrological models. GRACE has a limited temporal and spatial sampling capability, which results in aliasing errors, caused by high frequency processes. Although these processes are partly corrected for by means of background models of ocean and atmosphere, a great amount of noise persists. Furthermore, the free orbit geometry of

GRACE results in an uneven distribution of the ground tracks in time. Spatial visual-
isation of these ground tracks together with the rainfall, occurring during the month,
demonstrate that in some months, GRACE measurements of a certain target area, are
concentrated in a certain part of the month. This suggests that the time stamp of the
monthly solution for this area should be more weighted towards the part of the month
wherein most sampling is performed. Spatial filters, used to decrease the effect of alias-
ing errors, cause 'leakage errors'. It is shown that leakage errors considerably reduce the
information content of GRACE due to a decrease in both the yearly amplitude and the
interannual variability of storage. In order to use GRACE for hydrological modelling, an
iterative modelling approach is proposed that jointly accounts for the correction of leakage
errors and the validation of a chosen hydrological model structure as hypothesis for the
dominant hydrological processes. This method has been successfully applied in the upper
Zambezi and its surroundings in Chapter 5.

In Chapter 6, a new framework is proposed for the calibration of models with a combi-
nation of scarce, intermittent and non-concomitant data, and complementary data that
contain soft information. The method allows for the use of a multitude of signatures in
long-term data as explanatory information for calibration, instead of the usual residual
time series of an observed and modelled signal. These signatures are introduced in the
calibration as limits of acceptability to establish whether or not a parameter set is fea-
sible, given the information content. The limits are derived from the data as objectively
as possible, by using the year-to-year variability of the signature. Furthermore, limits of
acceptability for soft information, for instance derived from satellite observations, may be
introduced based on the perception of the modeller of the effective uncertainty of the soft
information.

The calibration method has been applied on a model of a subcatchment of the Zambezi,
the Luangwa. Here, only satellite rainfall and scarce, with respect to the rainfall non-
concomitant, discharge data are available for calibration. In Chapter 7, a time series of
evaporation maps has been introduced as soft information to further reduce the feasible
parameter space and to spatially distribute the parameters. Finally, the model with the
resulting constrained parameter space has been validated on one year of data, collected
during a field campaign, consisting of a number of rain gauges and water levels.

The development and application of methods to integrate satellite information in hydro-
logical models has led to the following important conclusions:

- In order to optimally use GRACE observations in the construction of hydrological
  models, a cooperation is needed between the work fields of geodesy and hydrology.
  The geodesist offers optimally filtered GRACE solutions. In turn, the hydrologist

may deliver an estimate of leakage errors. This can be done by encapsulating the modeller's prior knowledge of the hydrology in a hydrological model of the basin and its surroundings up to an extent, where the effect of the filter becomes negligible. The effect of the filter, applied by the geodesist, may be estimated from the storage estimates, delivered by this first modelling attempt. This enables a correction of the GRACE solutions for leakage errors. The corrected solutions may in turn be used by the hydrologist to validate or improve the hydrological model.

- The case study in the upper Zambezi showed that the combination of ground data, GRACE and expert knowledge of the hydrology of the study area, leads to a robust model structure. The expert knowledge ensures that the search space of model structures is beforehand significantly reduced. In this respect, a field visit is of great importance to collect this prior knowledge.

- In order to effectively use scarce, uncertain and non-concomitant data, the modeller has to be prepared to move away from one-size-fits-all objective functions. In these cases, a multitude of innovative limits of acceptability are necessary, which individually have a limited capability to constrain model parameters, but do have this capability when combined. It is up to the modeller to find criteria that exhibit a low sensitivity for the poor quality of the data.

- It is essential that the previously mentioned limits of acceptability are as much as possible estimated in an objective manner from available data. A cascading of subjectively chosen limits can result in considerably different results of the constrained parameter space.

- Remotely sensed evaporation as soft information, is capable of constraining model parameters related to the water balance of the soil moisture. Consequently, these data can be employed to spatially distribute these parameters.

- The combination of hard and soft limits of acceptability, based on various orthogonal data, enables a modeller to perform indirect calibration of hydrological models. This is considered to be proven by validation with an overlapping time series of rainfall observation (combined with satellite rainfall) and water levels.

Hessel Winsemius

Delft, July 2009

Samenvatting

Veel stroomgebieden in de wereld worden schaars bemeten. De weinige metingen die er wèl zijn in deze gebieden zijn vaak onzeker, schaars, onvolledig of ongelijktijdig. Dit bemoeilijkt de constructie van hydrologische modellen, die nodig zijn om op korte termijn droogtes, hoogwaters en beschikbaarheid van water te voorspellen, en op lange termijn de effecten van veranderingen in het landoppervlak of van het klimaat op de waterbalans en rivierafvoeren in te schatten. Gelukkig worden steeds meer gegevens, die mogelijk gebruikt kunnen worden in de ontwikkeling, kalibratie en validatie van hydrologische modellen, gratis aan het publiek beschikbaar gesteld. Met name satellietgegevens zijn hierbij zeer interessant. Belangrijke variabelen zoals regen, verdamping, straling, bodemvocht en waterberging kunnen tegenwoordig worden geschat op basis van ruwe satellietmetingen.

Dit proefschrift beschrijft de ontwikkeling en toepassing van methoden om de schaarse gegevens, die beschikbaar zijn in slecht bemeten stroomgebieden, te combineren met expertkennis en moderne satellietgegevens voor de conceptualisering, kalibratie en validatie van hydrologische modellen. De satellietinformatie wordt hierin gezien als 'orthogonale' informatie in de zin dat het een onafhankelijke kijk biedt op het functioneren van een stroomgebied. De methoden zijn zo ontwikkeld dat optimaal gebruik wordt gemaakt van de informatie, die vastligt in alle observaties, terwijl op een zo objectief mogelijke wijze rekening gehouden wordt met de onzekerheden in zowel grondmetingen als de gebruikte satellietgegevens. Er wordt aangetoond dat op deze wijze zelfs gegevens van slechte kwaliteit als modelinvoer en -uitvoer kunnen dienen. De gebruikte, complementaire satellietgegevens zijn satellietgebaseerde regenschattingen, gravitatiemetingen van de satellietmissie Gravity Recovery And Climate Experiment (GRACE) en satellietgebaseerde verdampingsschattingen. De methoden zijn ontwikkeld op basis van een aantal casussen in het stroomgebied van de Zambezi. Dit is een schaars bemeten stroomgebied, dat een uitstekende kans biedt om direct de ontwikkelde methoden in de praktijk te brengen.

In Hoofdstuk 4 wordt een analyse gegeven van de fouten in de GRACE observaties en hoe deze mogelijkerwijs het hydrologische bergingsignaal en daarmee de toepassing in de ontwikkeling van hydrologische modellen zouden kunnen beïnvloeden. GRACE heeft een beperkte temporele en ruimtelijke meetcapaciteit, wat resulteert in aliasing fouten,

veroorzaakt door hoog-frequente processen. Ondanks het feit dat er door middels van achtergrondmodellen van oceaan en atmospheer deels gecorrigeerd wordt voor deze processen, blijft er een grote hoeveelheid ruis bestaan. Verder resulteert de vrije loopbaan van GRACE in een onevenwichtige ruimtelijke verdeling van de satellietbanen in de tijd. Ruimtelijke visualisatie van deze satellietbanen, samen met de regen, gevallen binnen de maand, toont aan dat in sommige maanden GRACE metingen over een gebied zich concentreren in een bepaald deel van de maand. Dit suggereert dat het tijdstip van de maandelijkse oplossing voor dit gebied meer gewogen moet zijn naar het deel van de maand waarin metingen zich concentreren. Ruimtelijke filters, die toegepast worden om het effect van aliasing fouten te verminderen, veroorzaken 'lekkagefouten'. Er wordt aangetoond, dat lekkagefouten de informatie-inhoud van GRACE aanzienlijk reduceren door een afname in zowel de jaarlijkse amplitude als de jaar tot jaar variabiliteit van berging. Om GRACE te kunnen gebruiken ten behoeve van hydrologisch modelleren, wordt een iteratieve modelleermethode voorgesteld, die voorziet in zowel de correctie van lekkagefouten als de validatie van een gekozen modelstructuur als hypothese voor de dominante hydrologische processen. Deze methode is in Hoofdstuk 5 met succes toegepast in de Boven-Zambezi en de omgeving hiervan.

In Hoofdstuk 6 wordt een nieuwe methode voorgesteld voor kalibratie van modellen met een combinatie van schaarse, gebrekkige, ongelijktijdig gemeten gegevens en complementaire gegevens, die zachte informatie bevatten. De methode voorziet in het gebruik van een veelvoud van signaturen in langjarige gegevens als verklarende informatie voor de kalibratie, in plaats van de gebruikelijke tijdserie van residuen van een gemeten en gemodelleerd signaal. Deze signaturen worden in de kalibratie geïntroduceerd als acceptatielimieten om op basis van de informatie-inhoud vast te stellen of een parameterset reëel is of niet. De limieten worden op een zo objectief mogelijke manier uit de gegevens geschat door het gebruik van de jaar tot jaar variabiliteit van de signaturen. Verder kunnen acceptatielimieten voor zachte informatie, zoals bijvoorbeeld afgeleid uit satellietobservaties, worden geïntroduceerd op basis van de perceptie van de modelleur over de effectieve onzekerheid van de zachte informatie.

De kalibratiemethode is toegepast op een model van een deelstroomgebied van de Zambezi, de Luangwa. Hier zijn alleen satellietregen en gebrekkige, ten opzichte van de regen ongelijktijdige afvoergegevens beschikbaar voor kalibratie. In Hoofdstuk 7 is vervolgens een tijdserie met verdampingskaarten geïntroduceerd als zachte informatie om de reële parameterruimte verder te reduceren en parameters ruimtelijk te distribueren. Tenslotte is het model met de resulterende ingeperkte parameterruimte gevalideerd met een jaar gegevens die verzameld zijn tijdens een veldcampagne, bestaande uit tijdseries van een aantal regenstations en waterstanden.

De ontwikkeling en toepassing van methodieken om satellietinformatie te integreren in hydrologische modellen heeft geleid tot de volgende belangrijke conclusies:

- Om GRACE observaties optimaal te kunnen gebruiken in het construeren van hydrologische modellen is een samenwerking tussen de werkvelden van geodesie en hydrologie nodig. De geodeet biedt optimaal gefilterde GRACE oplossingen. De hydroloog kan op zijn beurt een schatting van lekkagefouten aanbieden. Dit kan hij doen door zijn voorkennis van de hydrologie in te kleden in een hydrologisch model van het stroomgebied en de omgeving hiervan, tot het punt waar het effect van het gebruikte filter verwaarloosbaar wordt. Het effect van het filter, toegepast door de geodeet, kan geschat worden op basis van de bergingschattingen, geleverd door deze eerste modelleerpoging. Hiermee kunnen de GRACE oplossingen gecorrigeerd worden voor lekkagefouten. Deze gecorrigeerde oplossingen kunnen tenslotte door de hydroloog gebruikt worden om het hydrologische model te valideren of te verbeteren.

- Uit de casus in de Boven-Zambezi blijkt dat de combinatie van grondgegevens, GRACE en expertkennis van de hydrologie van het studiegebied, leidt tot een robuuste model structuur. De expertkennis zorgt er hierbij voor, dat de zoekruimte van modelstructuren van tevoren al aanzienlijk gereduceerd wordt. Een bezoek aan het veld is daarom van groot belang om deze voorkennis te vergaren.

- Om effectief gebruik te kunnen maken van schaarse, onzekere en niet overlappende gegevens moet de modelleur bereid zijn om afstand te doen van one-size-fits-all doelfuncties. In deze gevallen is een veelvoud aan vindingrijke acceptatielimieten nodig, die individueel een beperkte mogelijkheid hebben modelparameters in te perken, maar gecombineerd wèl hiertoe in staat zijn. Het is hierbij aan de modelleur om criteria te vinden, die een lage gevoeligheid vertonen voor de matige kwaliteit van de gegevens.

- Het is essentieel dat voorgenoemde acceptatielimieten zover mogelijk op een objectieve manier geschat worden uit beschikbare gegevens. Een cascade van subjectief gekozen limieten kan resulteren in uiteenlopende resultaten van de ingeperkte parameterruimte.

- Remotely sensed verdamping als zachte informatie, is in staat om modelparameters in te perken, die gerelateerd zijn aan de waterbalans van het bodemvocht. Deze gegevens kunnen vervolgens ingezet worden om deze parameters ruimtelijk te distribueren.

- De combinatie van harde en zachte acceptatielimieten, gebaseerd op verschillende orthogonale gegevens, stelt een modelleur in staat om indirecte kalibratie van hydrologische modellen uit te voeren. Dit wordt bewezen geacht door validatie met

een overlappende tijdserie van regenmetingen (gecombineerd met satellietregen) en
waterstanden.

Hessel Winsemius

Delft, juli 2009

List of symbols

a	acceleration (L T^{-2})
a	calibration parameter, in Chapter 3 (-)
a	rating constant, in Chapters 5 and 7 (L^{3-b} T^{-1})
b	calibration parameter, in Chapter 3 (-)
b	rating constant, in Chapters 5 and 7 (-)
c	scaling constant, in Chapter 7 (-)
c_p	specific heat of air (L^2 T^{-2} K^{-1})
e_a	actual vapour pressure (M L^{-1} T^{-2})
e_s	saturated vapour pressure (M L^{-1} T^{-2})
f	spatial filter, in Chapters 4 and 5 (-)
f	spectral density, in Chapter 6 (-)
g	gravitational acceleration, 9.81 m s^{-2} (L T^{-2})
g_{ah}	aerodynamic conductance for heat transfer (L T^{-1})
g_s	surface conductance (L T^{-1})
h	water level (L)
i	slope (-)
l_p	soil moisture fraction where stress occurs (-)
m	mass (M)
n	number of precipitation stations (-)
p	pressure (M L^{-1} T^{-2})
q	specific humidity (-)
r_c	recharge coefficient (-)
r_{perc}	percolation rate (L T^{-1})
s	derivative of vapour pressure curve over temperature (M L^{-1} T^{-2} K^{-1})
u	wind velocity (L T^{-1})
$u_\star$	friction velocity (L T^{-1})
w	specific humidity, in Chapter 3 (-)
z	height above surface (L)
z_{0m}	roughness height (L)
A	surface area (L^2)
B	runoff generation power shape (-)
D	interception threshold (L T^{-1})
E	evaporation (L T^{-1})
E_i	evaporation from interception (L T^{-1})
E_t	transpiration (L T^{-1})
E_p	potential evaporation (L T^{-1})

E_{tp}	potential transpiration (L T^{-1})
F_{perc}	maximum percolation rate (L T^{-1})
G	ground flux (M T^{-3})
H	sensible heat flux (M T^{-3})
J	estimate of spectral density (-)
K_q	quick reservoir reciprocal of residence time (T^{-1})
K_s	slow reservoir reciprocal of residence time (T^{-1})
L	likelihood function, Chapter 6 (-)
M	model (-)
P	precipitation (L T^{-1})
P_g	ground based precipitation (L T^{-1})
P_m	merged precipitation (L T^{-1})
P_n	net precipitation (L T^{-1})
P_s	satellite based precipitation (L T^{-1})
Q	discharge (L^3 T^{-1}/L T^{-1})
Q_q	fast discharge (L^3 T^{-1}/L T^{-1})
Q_s	base flow (L^3 T^{-1}/L T^{-1})
R	radiation (M T^{-3})
S	terrestrial water storage (L)
S_{max}	maximum root zone water holding capacity (L)
S_q	quickly released storage (L)
S_s	slowly released storage (L)
S_u	soil moisture storage (L)
T	temperature (K)
$T_\star$	temperature scalar (K)
V	volume (L^3)
γ	psychrometric constant (M L^{-1} T^{-2} K^{-1})
η	scaling factor (-)
θ	parameter set (-)
κ	von Karman's constant (-)
λ	latent heat of vaporization (L^2 T^{-2})
ρ	density (M L^{-3})
ϱ	universal gas constant (M^2 L^2 T^{-2} K^{-1})
τ	atmospheric transmissivity (-)
ψ_m	stability correction for momentum (-)
ψ_h	stability correction for heat (-)
ω	fourier frequency (-)

Λ Obukhov length (L)
Φ autoregressive model

Contents

Chapter 1

Introduction

1.1 Hydrological modelling: bottom-up or top-down?

Over the last decades, the increasingly available computational power has led to the development of increasingly complex and distributed models, requiring a large amount of input and parameter calibration. Often, the justification of such models lies in their application: studying spatially highly variable and complex problems such as land cover change, erosion, contaminant transport, anthropogenic impacts and climate change. Abbott *et al.* (1986a) mentioned: "only through the use of models which have a physical basis and allow for spatial variations within a catchment can these problems be tackled", which was his justification for the development of the SHE model (Abbott *et al.*, 1986a, 1986b). Only years after the development and after numerous calibration and application efforts were the large shortcomings of this approach recognised. The input requirements were enormous and the calibration of model parameters led to an infinite amount of possible model parameter combinations that basically reproduced the same signal, which in general was an observed discharge hydrograph. Beven (1989) was one of the leading hydrologists who highlighted the problems related to dimensionality, representativeness of scale and false applicability of the by that time state-of-the-art physically based models. The term *equifinality* was adopted from systems theory to describe the effects of the high dimensionality (i.e. too many degrees of freedom in contrast with sparse evidential data) of most physically based distributed hydrological models on predictive uncertainty. Other elaborate efforts in physically based modelling at the representative catchment scale, in the hope that a model can be found that does not require calibration, were performed by Reggiani *et al.* (1998); Lee *et al.* (2005); Zhang and Savenije (2005). Nonetheless, for generic catchment-scale problems, the use of *parsimonious* (conceptual) models was and often is still advocated (e.g. Beven & Binley, 1992; Beven & Freer, 2001; Savenije, 2001).

Physically based models are examples of a *bottom-up* modelling approach. In hydrology, bottom-up modelling means that emerging small-scale behaviour (for instance measured in

a laboratory) is expected to remain a dominant feature at the large scale (for instance the scale of a catchment). This expectation is in general almost never met, because at larger scales, other properties than measured at the lab scale become dominant. An example of such "scale breaks" is the conductivity of a soil sample, measured in a laboratory. If we consider that the sampled conductivity is representative for the entire catchment, even if we take the average of multiple measurements, then we ignore the fact that within the soil matrix, there are root channels, animal burrows, cracks and fissures. This is a classical example of the problem of scale breaks and scale representativeness. The scale dependence of dominant runoff generating processes has only recently become subject of study (Didszun & Uhlenbrook, 2008).

Conceptual models (rather than physically based models) can be used to understand dominant processes and time scales in a catchment and to quantify them. In such research, a hypothesis is formulated that is continuously subjected to falsification efforts by introducing information. This information can either be retrieved from expert knowledge or from any new observation that comes available. The information acts as a guide, both for the formulation of the model structure (i.e. what are the dominant processes, how are they related to each other) and for quantification of its parameters. This is a *top-down* approach. Rather than being based on physical equations, these hypotheses are generally based on physical knowledge of the catchment's response on rainfall and are coded in the form of simple mathematical relations which can be easily adapted if the (new) evidential information conflicts with the prior assumptions made (e.g. Fenicia *et al.*, 2008b). A classical example of falsification has been presented by Sklash and Farvolden (1979), who wrote one of the first papers where the established hypothesis about insignificance of groundwater contributions to stream flow during storm periods was falsified with measurements of concentrations of environmental isotopes. This resulted in a large amount of studies in which attempts were made to separate stream flow contributions with environmental and other tracers (e.g. Uhlenbrook & Leibundgut, 2002; Fenicia *et al.*, 2008a). It provided justification for more complex, process-based models. Another interesting way of further reducing predictive uncertainty is by including expert knowledge about dominant processes and their magnitude. Expert knowledge is often an underrated source of information as will be shown in a further chapter of this thesis. Its value is perhaps not quantifiable, but it is in fact, an observation just like any other. Expert knowledge can easily be tested by introduction of fuzzy measures in the optimisation process (Seibert & McDonnell, 2002). Finally, remote sensing platforms offer the ability to indirectly observe rainfall, variations in terrestrial water storage, energy budgets and the energy balance. These indirect data sources can also be employed to constrain model uncertainty.
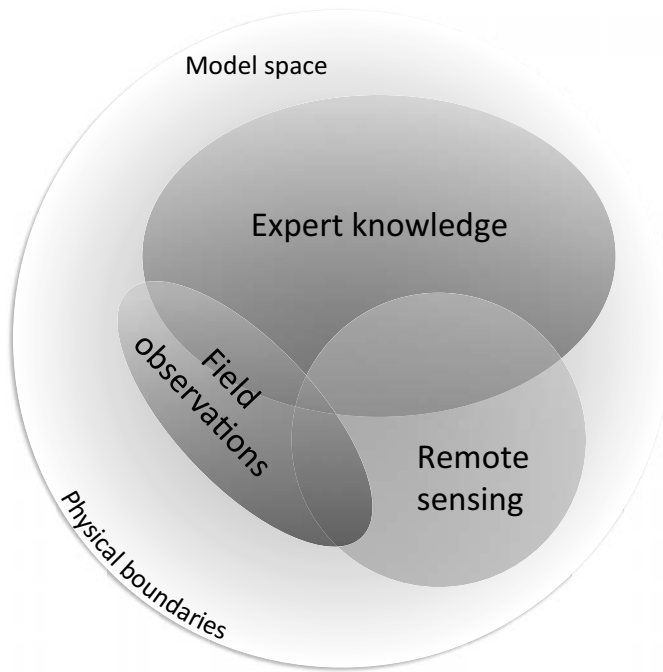
Figure 1.1: A visual sketch of the effect of multiple constraints on the model space as a whole. The model space is only constrained by physical boundaries, such as the amount of available water and energy. Theoretically, if all information available is combined in a multiple constraining framework, the predictive uncertainty will reduce to the area where all information converges. Information does not necessarily have to be hard, it can also be derived from expert knowledge or remotely sensed data.

1.2 Information as a means to constrain the model space

It is clear that the introduction of all kinds of information from independent observations, whether in the form of hard information, soft information (e.g. from satellite remote sensing) or expert knowledge, allows us to (sequentially) constrain the *model space* (Fenicia *et al.*, 2008a). The definition of 'hard information' is in this thesis, information of which the uncertainty or variability can be estimated in an objective manner, from the data available. Soft information on the other hand, can be of a more qualitative nature or can be information, of which the uncertainty cannot be straightforwardly demonstrated from the data*. "Independent" in this thesis, means that not the measurement itself is independent (e.g. a measurement may have been taken by a different instrument or at a different place) but measured processes and quantities are different. Such observations may provide a completely new, *orthogonal* insight in hydrological behaviour. Fig. 1.1 visu-

———

*.   note that in information theory, soft information is often referred to as information that poses a fuzzy constraint on a model

alises this process. Consider that the model space of a catchment is in principle infinitely large (except for the fact that physical principles such as the water balance and energy balance should be met). This means that any model structure with any parameter set can reflect the truth given that we do not have any information on the catchment. Now we start walking around in the catchment and we make certain observations. For instance: we observe that the governing climate consists of one wet season and we observe that each year, stream flow remains essentially zero in the first month of the rainy season. The fact that we know this, means that we can start rejecting all models within the model space that do not obey this fact. The result may be that we know that the 'truth' is somewhere within the large green circle in Fig. 1.1. This is an example of including expert knowledge in the modelling process. Of course the rejection criterium is rather subjective but when no other information is available, these constraints can be very valuable. Another example: now we do a questionnaire in the catchment and ask people whether they ever observed any water flowing over the land surface. In many catchments, the conclusion will probably be that this does not occur, which means that we can exclude Hortonian overland flow as a dominant process in the model structure. A last example: assume we have been investigating this catchment for about a year and we have concluded that the water in the streams has almost no turbidity, even during the rainy season when there is far more stream flow than in the dry season. This is an indication that most of the stream flow actually originates from groundwater. This puts another constraint on Fig. 1.1. Of course we can also start doing measurements, but the type of investigations mentioned above may help to define what we should measure to improve our process understanding.

1.3   Hydrological modelling under ungauged conditions; justification of this study

Although these observations are useful, the formulation of hydrological models is jeopardised in many parts of the world by the absence or deterioration of measurement networks. The scarce ground data available in such areas result in highly uncertain input (i.e. no or little ground rainfall measurements); high and perhaps ill-quantifiable parameter uncertainty; model structural uncertainty; and as a result ill-quantifiable predictive uncertainty of the outputs due to all these uncertainties. Ironically, these are exactly the areas were the need for hydrological predictions is high. This requires some further explanation: many ungauged regions are in Sub-Saharan developing countries. These are mostly semi-arid river basins (i.e. the energy availability for evaporation is much higher than the annual rainfall). The Budyko (1974) curve (Fig. 1.2) offers an opportunity to visualise the water balance in these regions. Semi-arid areas are on the 'moisture constrained' side of
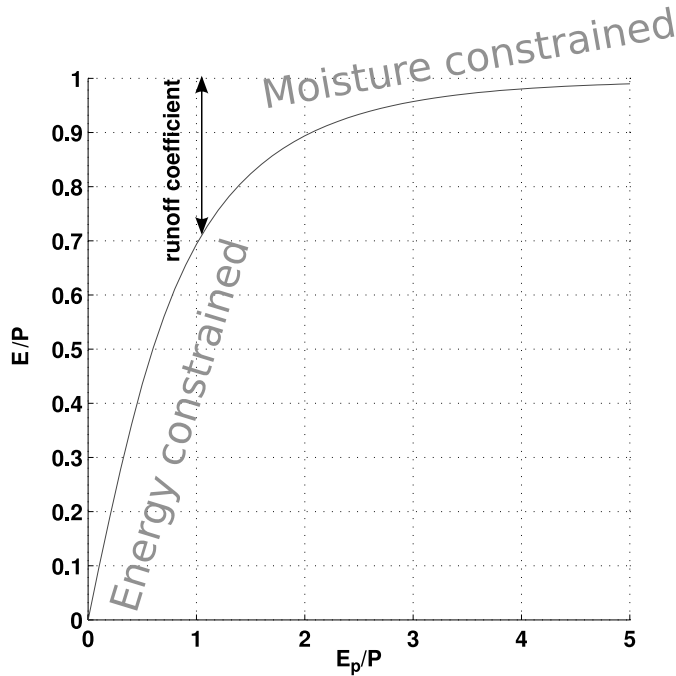
Figure 1.2: The Budyko curve. The graph shows the expected relation between an aridity index (E_p/P) and the water balance, i.e. how much of the rainfall evaporates (E/P). E_p is the annual potential evaporation, P is annual rainfall and E is actual evaporation. E_p/P is above 1 in semi-arid areas, which means that the evaporation is moisture constrained (i.e. limited by the amount of available moisture). In temperate climates, such as in large parts of Europe, the evaporation is energy constrained (limited by the amount of available energy).

the curve, which means that most of the rainfall will evaporate (i.e. the aridity index, $E_p/P > 1$). Discharge is only a small part of the water balance, which in practice means that when a hydrological year brings less than average rainfall, the discharge will decline at a higher rate than the rainfall. The opposite is true for wet years, where a small positive anomaly in rainfall can bring a flood. It also means that a modeller is challenged by the fact that a small error in available rainfall information, may result in a large uncertainty in the runoff response of his model. In calibration mode, this may lead to biased parameters and in predictive mode in considerable predictive uncertainty. However, these problems should not necessarily refrain hydrologists and water resource managers from further investigation. In first principle, this opinion is based on the fact that there is a great need for hydrological models in these ungauged basins, both in view of envisaged future changes in climate or land cover or land use. Second, often there is a lot of information available, but we just have to learn how to incorporate it in our models.

Initiatives to advance predictions under ungauged circumstances have been reported under the umbrella of Predictions in Ungauged Basins (PUB Sivapalan, 2003; Wagener *et al.*, 2004). Since then, work has been done on new measurement techniques (e.g. Uhlenbrook & Wenninger, 2006; Selker *et al.*, 2006; Westhoff *et al.*, 2007), new modelling frameworks (e.g. Reggiani *et al.*, 1998; Zhang *et al.*, 2006; Lee *et al.*, 2007; Schymanski *et al.*, 2008), use of new and soft data sources (e.g. Seibert & McDonnell, 2002; Franks, 2007; Immerzeel & Droogers, 2008), new optimisation techniques (e.g. Vrugt *et al.*, 2003; Kuczera *et al.*, 2006), optimal use of available data (e.g. Atkinson *et al.*, 2002; Wagener, 2003; Montanari & Toth, 2007) and improvement of model structures by means of multi-informative optimisation (Vaché & McDonnell, 2006; Son & Sivapalan, 2007; Fenicia *et al.*, 2008a). All these contributions may move forward our ability to predict streamflow in ungauged basins.

Most PUB-related studies that use some form of complementary ground data, demonstrate modelling approaches in small or meso-scale catchments. The type of ground data used for model improvement include, besides streamflow, complementary groundwater levels or tracers used for hydrograph separation, end member analysis or mean residence time estimation. The use of such measurements becomes problematic at the larger scale, where dense spatial sampling becomes infeasible and the integration of discharge contributions from various subcatchments results in averaging of the observed tracer signatures. Moreover, at large spatial scales, lumping of all responses cannot be justified, simply because of heterogeneity in hydro-climatology and/or properties of the land (sub)surface. The result is that models at these scales have to be simplified to parsimonious model structures in order to prevent too many degrees of freedom given the data available, or their over-parameterisation has to be accepted in order to fulfil the objective of these models. One of the few information sources that can provide constraints to hydrological models of ungauged basins at these scales is satellite remote sensing.

1.4   Problem definition and study objectives

In ungauged basins, hydrological modelling at large scale has to deal with limited available data. This means that classical calibration, routinely applied in gauged basins, has to rely on precipitation records and hydrometric records that are not available, of poor quality, sparsely distributed in space or non-concomitant in time. Furthermore, in semi-arid areas discharge is on average a small flux, not directly proportional to the average rainfall due to evaporation-related threshold processes. This aggravates uncertainties related to rainfall input uncertainty. As a result, relevant hydrological processes can often not be identified

from the sparse data available. Hence these processes are disregarded or conceptualised wrongly or as part of a different process. This results in biased and weakly identifiable model parameters and high model structural and parameter uncertainty. Even though this may sometimes not be directly reflected in output uncertainty (e.g. because a model is still able to reproduce a variable of interest within reasonable limits), the robustness of such models for predictions of future change (i.e. man-made and/or climate) may be poor.

Promising research venues that may increase confidence in hydrological models in data sparse basins involve the use of satellite remote sensing information as complementary data to reduce model uncertainty. Unfortunately, the satellite state of the art does not include a mission that is specifically designed for hydrological purposes. As a result, hydrological variables that may be derived from space-based measurements are a result of (often complex) transformation models that convert the raw observations to hydrologically meaningful variables. This process introduces uncertainty in the data itself that is not easily quantifiable. Therefore this thesis aims at using satellite information as *complementary* information to the formulation of hydrological models, while considering historical ground information that may possibly be of poor quality due to measurement uncertainties, intermittence and non-concomitance, as a foundation.

The main objective of this research is therefore to develop frameworks for using satellite information to drive, constrain and validate hydrological models, complementary to the use of traditional (ground based) data sources. The specific objectives are:

1. to study the information content of space-based data sources. This involves the use of gravimetric information from the Gravity Recovery And Climate Experiment (GRACE), which can directly be related to changes in water storage, and thermal-infrared imagery, which relates to the energy balance of the land surface. In addition, satellite-based rainfall (serving as model input) and variables from atmospheric models are considered.

2. to formulate methods to deal with poor quality of available input and output records in the calibration of models. This poor quality may be due to measurement uncertainties, intermittence, non-concomitance and limited spatial and temporal resolution.

3. to demonstrate the combined objective use of multiple information sources: (a) for assisting in the formulation of conceptual model structures for ungauged basins; and (b) for reduction and objective quantification of model uncertainty. *Information* in this context can be ground data, remotely sensed data and expert knowledge.

All study objectives are addressed through a number of case studies within the Zam-

bezi river basin, where conceptual hydrological models are formulated with the aid of aforementioned information sources.

1.5   Thesis outline

The study area, motivation for its choice and description of hydrology and water resources is presented in Chapter 2. A description of the data sources used in this study and the way they are processed is given in Chapter 3. In the four following chapters, real cases are presented that show the integration of models with satellite remote sensing in more detail. Chapter 4 describes in detail the possible error sources of one of the prime remotely sensed data sources considered in this thesis, information from the Gravity Recovery And Climate Experiment (GRACE). It discusses how these error sources may affect the usefulness of GRACE for hydrological applications and an iterative framework is proposed which deals with some of the major error sources. In Chapter 5, the proposed framework is demonstrated by developing a monthly scale semi-distributed model of the upper Zambezi and its surroundings. The spatial variability of the hydrological information, coming from this model, assists the GRACE community in correcting GRACE storage estimates for the so-called leakage error, whereafter GRACE may be used as validation for the hydrological model itself. It implies that if validation is not satisfactory, the model structure may be adapted several times, and new estimates of leakage errors may be derived from it, in order to reach a satisfactory resemblance between GRACE and the hydrological model. Chapters 6 and 7 describe an elaborate case study of one of the major tributaries to the Zambezi, the Luangwa river. A multi-objective calibration framework is presented in Chapter 6, which may be used in typical ungauged circumstances, where data is scarce, intermittent, of poor quality and non-concomitant. The framework deals with subjectivity, inherent in classical calibration approaches, by imposing limits of acceptability on information signatures, identified from scarcely available data. The used signatures are relatively insensitive to the accuracy of the used rainfall inputs. In Chapter 7, land surface related parameters are further constrained and distributed in space by means of satellite-based evaporation estimates. A validation is performed using a concomitant time series of satellite rainfall, in-situ rainfall and in-situ hydrometric measurements of one rainy season. The thesis ends in an elaborate discussion and conclusions, summarised in Chapter 8.

Chapter 2

Study area: the Zambezi river basin

2.1 Why the Zambezi?

Arguments used against the selection of the Zambezi as a study area are in general related to problems of data availability. It is a truly ungauged catchment, which means that it is unavoidable that problems with data availability occur. The available data are usually scarce, incomplete, subject to unknown accuracy and non-concomitant. New (often satellite-based) data sources are emerging in a period where in this part of the world, in-situ measurement networks have deteriorated, which compromises their joint use due to non-concomitance (see for example Fig. 2.1).

At the same time, the Zambezi is a semi-arid river basin, which means that the ratio of annual discharge to rainfall is very small, in the order of 10% (see Fig. 1.2). The high energy availability results in evaporation being the most dominant flux, while only a small surplus of rainfall will be converted to discharge. This surplus however, will be significantly and disproportionally larger during extremely wet years and significantly smaller in dry years. This non-linearity also results in a large uncertainty in stream flow estimation when rainfall is uncertain.

The selection of the Zambezi is therefore not related to the ease of application of methods or to data availability. It was chosen for three reasons: first, the research on PUB needs a move from experimental basins to real-life basins. There is a lot of literature on studies performed in catchments with numerous measurements that may be used to test hypotheses, to construct model theories and to apply optimisation strategies (e.g Seibert & McDonnell, 2002; McGlynn *et al.*, 2002; Uhlenbrook *et al.*, 2002; Selker *et al.*, 2006; Fenicia *et al.*, 2008a). These catchments are generally classified in between headwaters and meso-scale catchments. In the author's view, the studies performed are extremely valuable for improving the understanding of catchment behaviour but do not provide a solid basis for applications in real-life problems which usually apply when dealing with
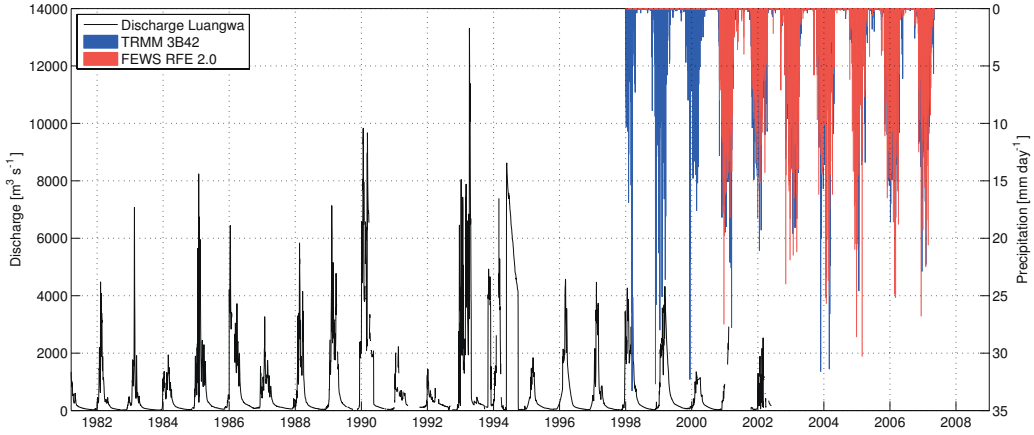
Figure 2.1: A typical example of data availability in the Zambezi river basin. The quality of the available discharge data deteriorates in more recent periods, while the benefit of having continuous near real-time daily rainfall estimates from space only starts in 1998. Note that local government agencies have also been collecting rainfall estimates before the onset of remotely sensed rainfall data sources, however these are often not readily available for researchers or water resources planners or they do not have the necessary spatial coverage.

much larger river basins. So far, the author of this thesis is not aware of a catchment of the order of magnitude of a tributary to the Zambezi, where intense measurement campaigns are performed, comparable to the nature of the ones, performed in the aforementioned studies. Such measurement campaigns are compromised due to a number of reasons. One of them is the size of the catchment: you simply cannot measure everything every square kilometre, and tracer signatures generally give poor system identification at these spatial scales. Other reasons are remoteness and sometimes even unwillingness of involved parties to share information at this scale (which, needless to say, was not the case during the realisation of this thesis). Therefore, we should find techniques that still allow us to reduce uncertainty about the current hydrological regime, given these restrictions. Under these circumstances, satellite remote sensing can offer great opportunities.

The second argument relates to the hydrological regime in the typical semi-arid climate of the Zambezi. The very fact that streamflow estimation may be subject to large uncertainty due to relatively small uncertainties in the rainfall inputs, is a challenge as such to address. To deal with these uncertainties we have to come up with robust methods to optimally use the information content of measurements of a small flux in the water balance (i.e. stream flow) and promising new satellite-based information sources that may help us to better quantify the water balance, while taking into account the aforementioned uncertainties.

This leads to the third argument for the selection of the Zambezi as a study area: the
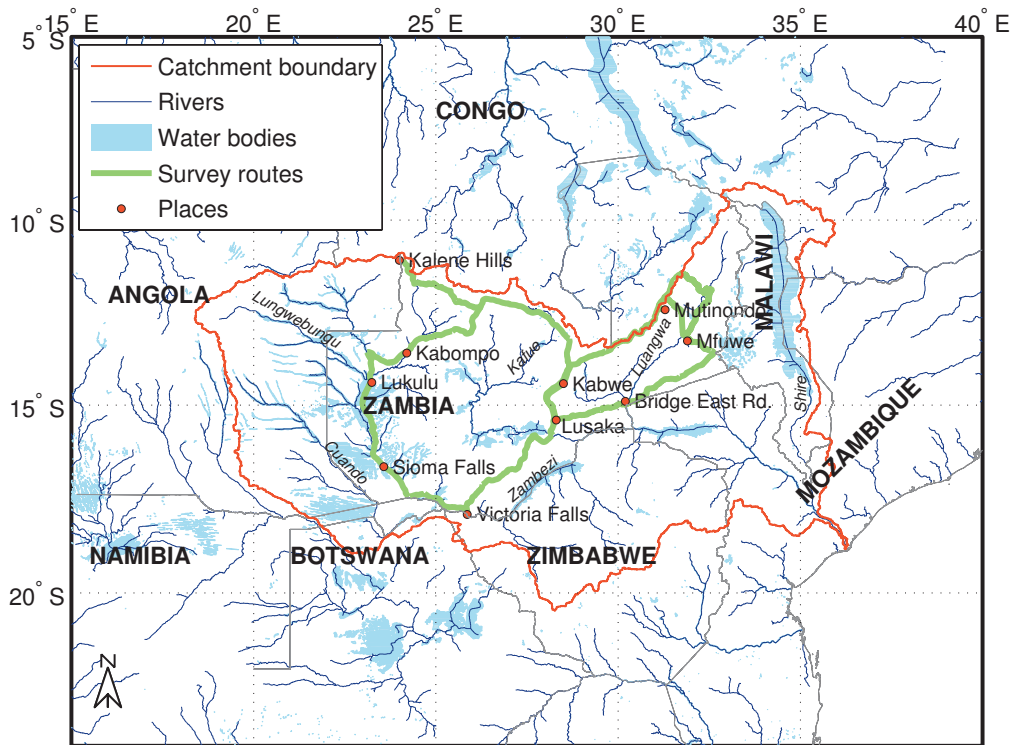
Figure 2.2: Zambezi basin, outlined in red. Blue indicates the major water bodies. Most of them are not permanent.

Zambezi is in particular an area where large variations in water storage are observed over the year. It is therefore a suitable environment to experiment with model improvements and constraints by means of GRACE satellite observations of water storage variations.

2.2 General information on hydro-climatology

The Zambezi (Fig. 2.2) springs from Zambia in the Kalene Hills, where it almost immediately flows into Congo. From there, it enters Angola, where a number of wetland dominated tributaries flow into the Zambezi. It flows back into Zambia again, where it merges with a number of large tributaries such as the Kabompo, flowing in from the East

and the Lungwebungu, flowing in from Angola in the West. The hydrology of these tributaries is dominated by floodplains and wetlands (some of the larger ones marked in blue in Fig. 2.2), locally called dambos that are able to store a considerable amount of water in the wet season and function as an interconnected system of fill and spill reservoirs. Downstream of Lukulu, the Zambezi flows into the Barotse floodplains, a seasonally inundated area, consisting of very deep unconfined Kalahari sand layers. The floodplains are about 20 km wide and 100 km long. Downstream of the floodplains, the Zambezi becomes the border of Namibia, Botswana, Zambia, Zimbabwe (where it drops down the Victoria falls and flows into Lake Kariba). North of Lake Kariba, the Kafue flows into the Zambezi. The Luangwa, one of the major tributaries, flows into the Zambezi, closely upstream of Lake Cahora Bassa, which represents the border between Zambia and Mozambique. The Shire originates from Malawi and is the major tributary in Mozambique.

(a)   (b)
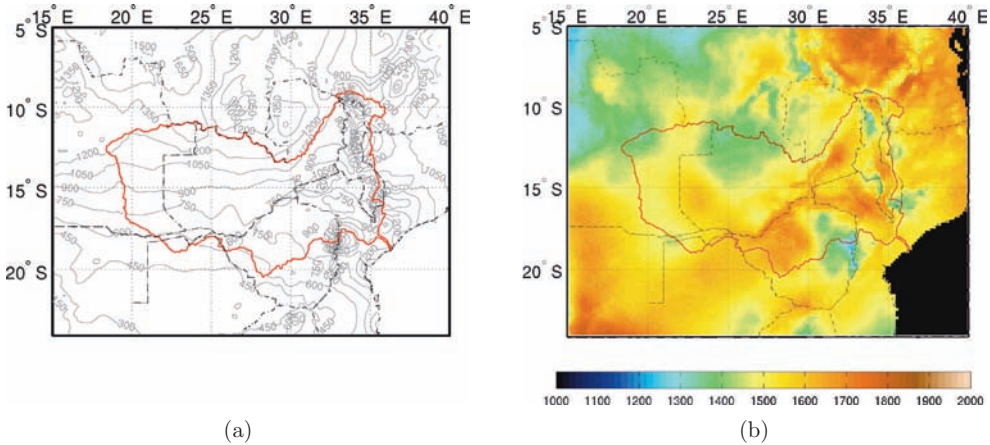
Figure 2.3: Rainfall (a) and reference evaporation (b) climatology of the Zambezi [mm yr^{-1}]. This climatology is based on the 1960-1990 meteorological climatology of the Climate Research Unit (CRU) and has been computed by the Food and Agricultural Organization (FAO, Allen *et al.*, 1998).

Rainfall in the Zambezi is about 600 mm yr^{-1} in the South and can be as much as 1200 mm yr^{-1} (based on New *et al.*, 2002) in the North (Fig. 2.3a). In extreme dry years, it can be as low as 400 mm yr^{-1}, but amounts up to 1500 mm yr^{-1} have also been recorded in the North. Rainfall is largely controlled by the Intertropical Convergence Zone (ITCZ), which moves over the Zambezi from October until April, causing concentration of rainfall in the months December until March. The vegetation generally follows the climate very well. In the end of the dry season, the areas receiving least rainfall and most potential evaporation have a very low Normalised Difference Vegetation Index (NDVI, a satellite-derived estimate for the greenness of the surface). This is clear from comparison of Figs. 2.4b and 2.3.

(a) NDVI in April, beginning of the dry season   (b) NDVI in October, end of the dry season
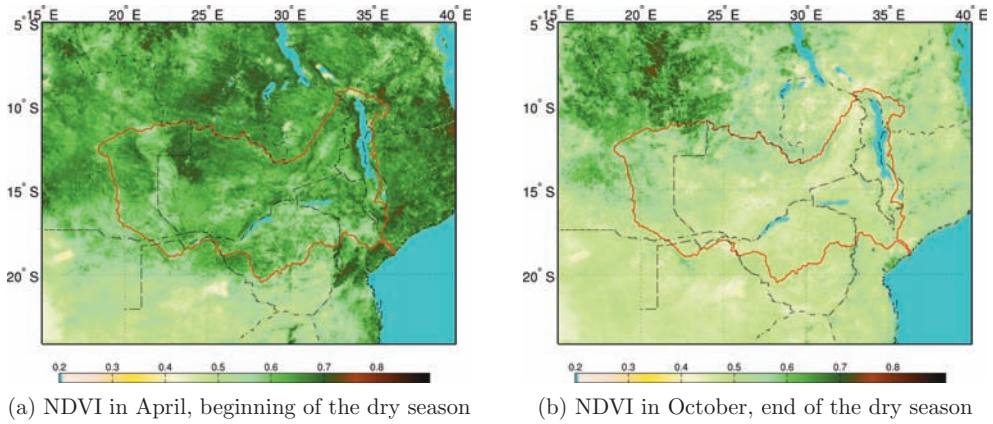
Figure 2.4: Difference between NDVI [-] in the beginning of the dry season and the end of the dry season.

The semi-arid character of the Zambezi is clearly revealed by the runoff coefficient, which can be estimated by applying the equation of Budyko (1974) on the precipitation and evaporation regime. This is given in Fig. 2.5. Runoff coefficients are ranging between 10 % in the South and 25 % in the North-West.

2.3   Water resources management

The Zambezi is shared by 8 riparian countries: Zambia, Angola, Namibia, Botswana, Zimbabwe, Tanzania, Malawi and Mozambique. Management of its water resources should therefore be done in an integrated manner, both quantitative and qualitative as to maximize the general value of water resources (see for an interesting approach to estimating the value of water Hoekstra *et al.*, 2001). Many agreements and development strategies have materialised under the Zambezi Action Plan Project 6 (ZACPRO 6) as part of the Zambezi Action Plan (ZACPLAN), which aims at integrated water resources management (IWRM) with regional cooperation. There are many stakeholders involved in the allocation of water resources, ranging from small-scale local users (e.g. agriculture and fisheries) to large agricultural companies, copper mining and hydropower. The many stakeholders, combined with the extremes related to the semi-arid nature of the river basin, makes ZACPRO a challenging program.

2.3.1 Some of the water resources users

Water resources of the Zambezi are allocated for drinking water, agriculture and industries, mainly copper and hydropower. In fact, mining of copper is one of the major growing
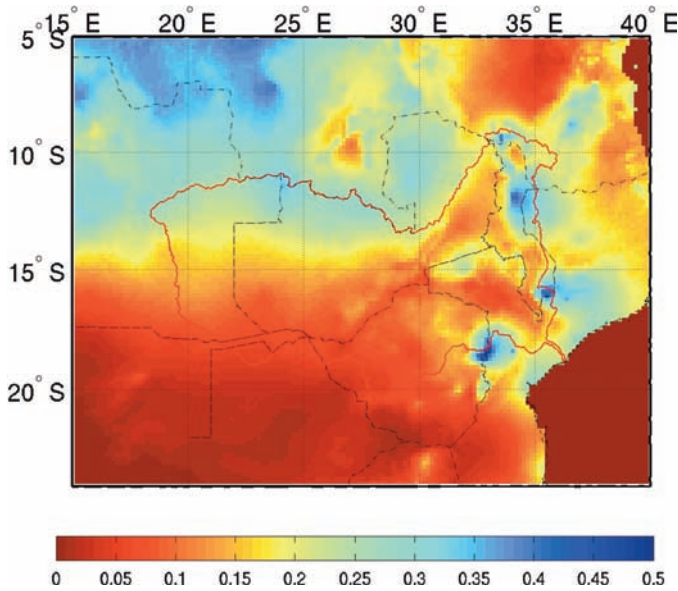
Figure 2.5: Runoff coefficient [-], estimated by application of the Budyko curve. Reference evaporation was taken as potential evaporation, from Fig. 2.3b.

industries at the moment and accounts for the major part of Zambia's economy, operating in the North and North-western part of Zambia, the Copperbelt region. The implications of this industry may jeopardise the water quality and thus reduce the downstream value of water, in particular in the Kabompo and Kafue tributaries in the upper Zambezi in Zambia. Von der Heyden and New (2004) showed for instance that in some of the many small impoundments where water is retained for mining purposes, contamination plumes in the groundwater can be observed, containing heavy metals such as cobalt, nickel and zinc. Luckily, in the studied areas, the soil has enough adsorption capacity to reduce the risk of drinking water contamination, but in other less researched areas with low adsorption capacity, for instance on crystalline bedrock, this may have serious implications for drinking and irrigation water quality and public health.

Another large industry having a great impact on water resources is the power industry. Lakes Kariba and Cahora Bassa belong to the largest man-made lakes in Africa, delivering power to parts of Zambia, Zimbabwe, Mozambique and South Africa. The first is operated by a bi-lateral organisation called the Zambezi River Authority (ZRA), the latter by Hydroélectrica de Cahora Bassa, which is fully owned by Mozambique since 2007. Other important reservoirs are located in the Kafue tributary and include Ithezi-thezi and Kafue Gorge. The controlled flow, especially in the lower Zambezi (downstream of Kariba) and

the Kafue river has seriously reduced the seasonal variability in inundation in the downstream areas, which has had serious ecological, social and economic implications (e.g. due to the reduction in shrimp catches, Gammelsrod, 1996). The annual returning floods in the Kafue flats are immensely important to sustain wildlife and fisheries. Since 1999 the World Wildlife Fund is supporting a dialogue between stakeholders of the Kafue catchment, which should result in restoration of the annual returning floods by a controlled flood release while maintaining energy supply.

2.3.2 Floods

Management of the Zambezi's water resources is sometimes seriously jeopardised by the occurrence of floods. The most recent severe floods occurred in 2007 and 2008 in Mozambique, which were caused by a combination of forced releases from Lake Kariba (see the press release of the Zambezi River Authority in Fig. 2.6) and excessive rainfall in the Luangwa, Shire and some smaller tributaries in Zimbabwe and Mozambique. Fig. 2.7 shows hydrographs of water levels at Mfuwe, upstream in the Luangwa river (measured during the realisation of this thesis by the author and others) and Zumbo (measured by DNA, Mozambique), which is at the confluence of the Luangwa and the Zambezi. With better knowledge of upstream releases and response to rainfall in ungauged basins, these floods could have been less severe for instance by allowing early and less abrupt releases in the large reservoirs, or warnings could have been issued earlier. Unfortunately, notification of flood releases to downstream reservoir operators is often subject to serious delays (Vaz, 2000) and many tributaries to the Zambezi are presently either ungauged or not gauged in real-time.

ZAMBEZI RIVER AUTHORITY

Press Release

OPENING OF LAKE KARIBA SPILLWAY GATES

MONDAY 11TH FEBRUARY, 2008

Due to the current heavy rainfall over the entire Zambezi basin and also the expected continued rainfall as the Department of National Metrological Services outlook, notice is hereby given that one spillway gate will be opened on Monday 11th February 2008, at 12.00 hours, however it should be noted that more gates may be opened as necessary without further notice.

Due to possible flooding down stream of the Kariba Dam, the general public and communities living along the banks are kindly advised to take this notice seriously to avoid any loss of life and damage to property.

The situation will be closely monitored.

For further enquiries please contact:

Eng. CFG Mukosa   Eng. S Mhlanga
Director (Water Resources &   Senior Manager (Projects & Dam Safety)
Environmental Management)   Zambezi River Authority
Zambezi River Authority   New Administrative Block
32, Cha Cha Cha Road   21 Lake Drive
Kariba House   **Kariba, Zimbabwe**
Lusaka, Zambia

Tel: +260 211 227970/3   **Tel:**   +263 61 2605
Fax: +260 211 227498   **Tel:**   +263 61 2606

Figure 2.6: Notification of opening of one of the spillway gates at lake Kariba.
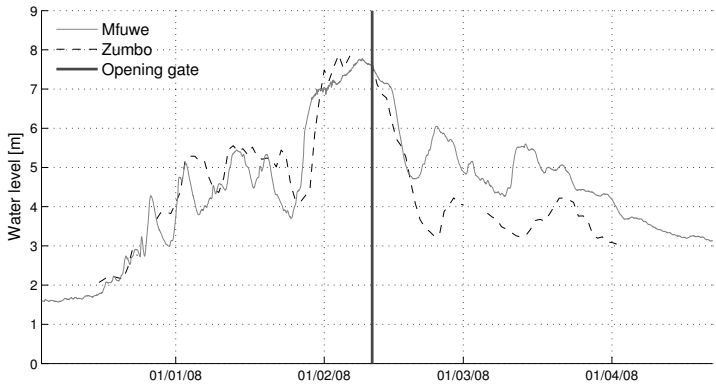
Figure 2.7: Hydrographs at Mfuwe (in the Luangwa tributary) and Zumbo (in the Zambezi), downstream Kariba, closely upstream Cahora Bassa. The moment of flood releases at Kariba, closely follows a flood peak in the Luangwa. This means that operators at Cahora Bassa had to deal with two simultaneous flood waves, entering the reservoir.

Chapter 3

Remote sensing data

The availability of data increases rapidly and data become more easily accessible as well. Even during the realisation of this thesis, many new data sources were introduced to users world wide through internet. Others are more experimental or classified and are only available through research centres.

This chapter gives a brief overview of the main characteristics of the most important data sources, used throughout this thesis, and the way in which they may be applied in hydrology. First, a simple explanation of the way the GRACE gravity mission measures the gravity field is given and its applicability in hydrology is briefly described. Second, a crucial input for hydrological models is rainfall. Some of the available techniques to derive rainfall estimates from satellite information, and some end user products are described. New promising information may be retrieved from geostationary satellites. An incoming short-wave radiation product, derived from Meteosat Second Generation images, used in parts of this thesis, is described and a validation is shown. Finally, techniques are shown how to use derivatives from atmospheric models. A physical downscaling algorithm, developed during an MSc. study, and further extended in this thesis is described and a validation is shown.

3.1   GRACE terrestrial water storage

The Gravity Recovery And Climate Experiment (GRACE) is a twin-satellite mission, developed to measure changes in the Earth's time-variable gravity field with unprecedented accuracy (e.g. Tapley & Reigber, 2001; Tapley *et al.*, 2004). The main mission objective of GRACE is to map global changes in mass due to the continental water cycle. With a minimum spatial resolution of about 200–300 km, it can be used to identify mass changes due to variations in water storage. This information can assist in determining groundwater depletion (e.g. Rodell & Famiglietti, 2002), ice melt (Velicogna & Wahr, 2006),

residual basin-scale estimates of evaporation (Rodell, Famiglietti, *et al.*, 2004; Ramillien *et al.*, 2006), multi-objective calibration of large-scale hydrological models (Güntner, 2008; Werth *et al.*, 2009) or validation of hydrological models (e.g. Klees *et al.*, 2007; Niu *et al.*, 2007; Ngo-Duc *et al.*, 2007; Rodell *et al.*, 2007; Tapley *et al.*, 2004; Wahr *et al.*, 2004; Hasan, 2009).

GRACE consists of 2 polar orbiting satellites, flying at an altitude ranging from 300 to 500 km, separated by a distance of about 200 km along track. The Earth's gravity field causes accelerations of the satellites, where they approach an area of relatively high mass concentration, and deceleration where they move away from them. The raw measurements consist of extremely accurate distances between the two satellites, measured by the High Accuracy Intersatellite Ranging System (HAIRS). The acceleration - deceleration behaviour of the two satellites can therefore be estimated and translated into estimates of along-track changes in the gravity potential, assuming that these changes occur due to changes in the upper layer of the Earth only. With only few measurements of gravity potential, the number of possible mass configurations can be numerous. In the extreme case, one instantaneous measurement could be the result of an infinite number of mass configurations. The more measurements taken, the lower the amount of possible mass configurations and the higher the spatial resolution that can be reached (one may compare it with GPS positioning, where measurements of only one GPS satellite give an infinite number of possible locations of a person in its surroundings). As a result, there is a trade-off between temporal and spatial resolution of GRACE, which is visualised in Fig. 3.1. Therefore, gravity fields are delivered, typically on a monthly basis, allowing the satellites to make several passes over each region of the Earth.

Since gravity can be described as anomalies of a spheric shape, the changes in the gravity field are expanded into spherical harmonics. These are comparable with Fourier series, often applied in time series analysis, whereby the first degree represents the lowest (in this case spatial) frequency of the anomaly and the highest degree represents the highest frequency and thereby the smallest spatial features. Given that the Earth's circumference is about 40 000 km, degree and order 1 has a wavelength of 40 000 km, degree and order 2 a wavelength of 20 000 km and so on. In practice, the measurability of spatial features decays with higher detail due to instrumental errors and high temporal variability in atmospheric or oceanic processes that cannot be compensated for by background models. Due to this reason, the first GRACE releases gave meaningful results only up to a degree and order 70 (although GRACE data centres derive up a much higher degree and order), which is equivalent to a spatial resolution of about 600 km. Nowadays, the data processing and noise reduction techniques have improved considerably so that the resolution can be pushed up to degree and order 120 (about 300 km).
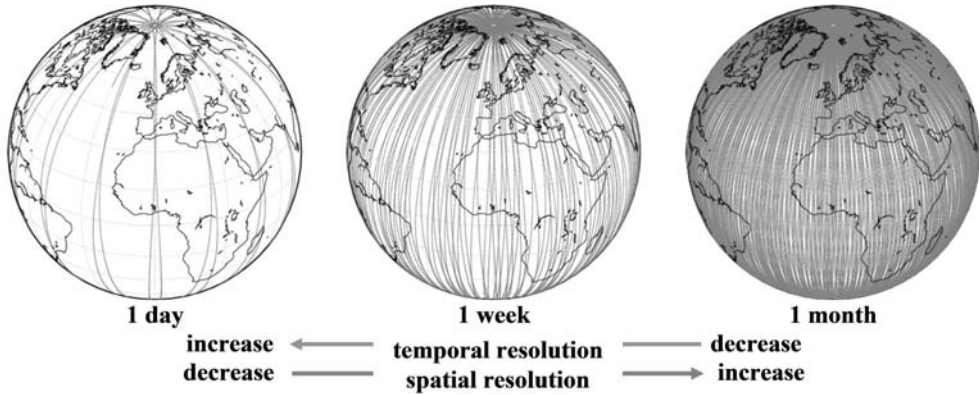
Figure 3.1: This figure shows from left to right how the spatial coverage of GRACE increases with time (1 day, 1 week and 1 month). It clarifies the trade-off between spatial and temporal resolution (Source, Schmidt *et al.*, 2008)

Changes in mass are caused by many processes, all occurring at different spatial and temporal scales. Some of the more important ones are gravitational pull by other mass bodies such as the Sun, Moon and nearby planets, atmospheric moisture redistribution, oceanic tides, but also deformation due to the latter process and post-glacial rebound. The high frequency processes, that are expected to vary a great deal within one month of data acquisition are corrected for by using several background models. The most important are oceanic models and atmospheric models. They are used to estimate the contribution of oceanic movements, Earth tides and atmosphere to the gravity solution, which is consequently subtracted from the total monthly mass anomaly. The residual gravity signal then represents unmodelled signals such as hydrology, tectonics, land deformation and noise, e.g. from instrumental errors and errors in the background models. The signal that is expected to vary most on a monthly time scale is the hydrological signal, which comprises terrestrial water storage changes that can be caused by variations in groundwater storage, soil moisture, snow pack and surface water (Fig. 3.2).

Note that the hydrological signal is assumed to be constant within the period of observation (i.e. one month) and it seems therefore logical to assume that for a given region, the time-averaged storage estimate is in the middle of the month. In Chapter 4, it is hypothesized that, depending on the size of the target area, this is often not the case and can be the cause of aliasing in the temporal behaviour of GRACE solutions. Uncorrelated spatial errors are usually filtered. A popular filter is a simple isotropic Gaussian smoother.

The possibilities of application of GRACE in hydrology are limited to the very large scale
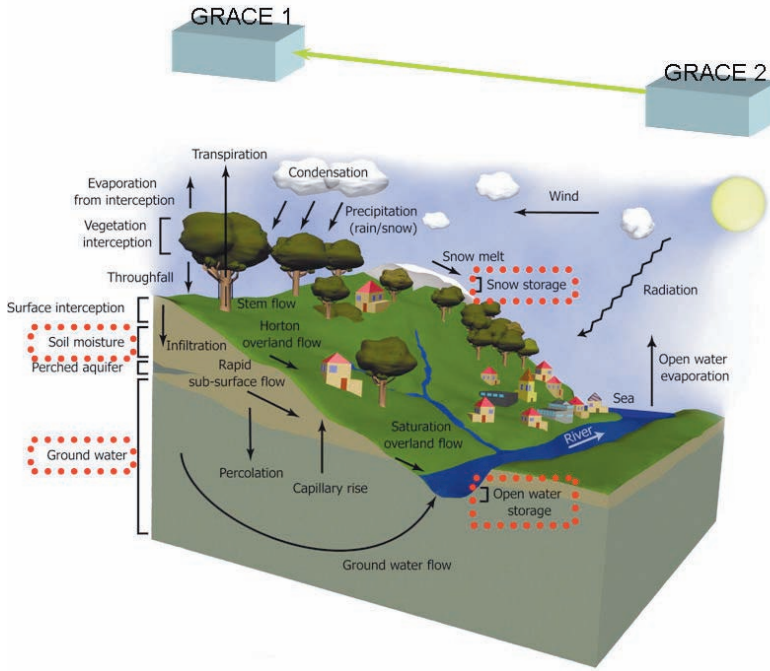
Figure 3.2: An artist impression of hydrological processes and storages. Storage, observed by GRACE only represents an integral of all compartments of water storage, appearing within the hydrological cycle. In this figure, the largest of these processes are encircled: soil moisture, groundwater, open surface water, snow and ice*

($\sim 200 \times 200$ (km)^2), which in my opinion is beyond the scale where one may assume spatial homogeneity. Throughout this thesis, it is therefore assumed that calibration efforts of hydrological models at such scales should be performed by means of other information sources than GRACE. The use of GRACE in model development lies in large-scale validation of model structures and parameter sets, which will be demonstrated in Chapter 5.

GRACE data are processed at three data centres: the Center for Space Research Texas (CSR) and the GeoForschungsZentrum Potsdam (GFZ) and the Jet Propulsion Laboratory (JPL). The Centre National d'Études Spatiales (CNES) and the Department of Earth Observation and Space Systems (DEOS), Delft, University of Technology, are creating their own solutions from the raw data (the gravity retrieval algorithm by DEOS is described by X. Liu, 2008). The differences in the solutions are due to different data retrieval and post-processing algorithms.

————————
*.   Acknowledgement to Miriam Gerrits for providing this figure.

3.2 Satellite-based rainfall estimation

3.2.1 Introduction

Rainfall is the driving input in water balance studies and hydrological modelling. Besides of course the ecological, geological and hydrological features of a river basin, it is the quantity and the temporal and spatial distribution of precipitation that determines how rainfall is partitioned over the three main fluxes in the water balance (evaporation, runoff and storage change). Especially in arid and semi-arid areas, the hydrological response in terms of river discharge behaves extremely non-linear with rainfall (see also Section 1.3), which means that small uncertainties in rainfall can lead to large uncertainties in discharge (in the extreme case it can mean either no discharge at all or, when system thresholds are exceeded, floods). Measurement networks have been set up to capture the space-time variability of rainfall, however in many areas in the world, precipitation monitoring networks are declining. In Africa, networks that were set up during colonial periods, are nowadays more and more abandoned due to lack of money or accessibility, the same also being the case with hydrometric measurement networks.

Fortunately, there are several new remote sensing platforms that allow us to make an estimate of the space-time variability of rainfall. It is important to note that satellite based rainfall estimates are in general not and probably will never be the sole solution to problems of data scarcity in ungauged and remote areas. Many satellite based rainfall estimates are made operationally but none of them describes the full features of the true rainfall at the land surface. Most algorithms are based on empirical relations between some variable, measurable from space, and ground rainfall. I say here 'ground rainfall' because 'observed' rain drops do not necessarily reach the ground. Until now, only one space borne precipitation radar has been launched on board the Tropical Rainfall Measuring Mission (TRMM), but even radar does not observe ground truth precipitation. The power of these satellite observations is revealed when merged with each other and with ground anchor points, where the assumed 'truth' is measured. In this manner, the different qualities of all estimates is optimally used (see Fig. 3.3 for an example of differences between estimates). Therefore I would like to stress that, although satellites may help a great deal in determining the space-time variability of rainfall, it is imperative that measurement networks are maintained and extended, especially at places where this variability is large. An opportunity for more extensive operational ground based measurements in sparsely gauged areas is tomography of rain rates through cell phone networks. Here the attenuation rate of signal power between two cell phone poles is power related to the rain rate (Messer *et al.*, 2006). For wind and temperature, cheap, smart sensors based on microchips (Makinwa & Huijsing, 2002; Partijs *et al.*, 2005) provide promising

technology to extend measurement networks. It would greatly benefit the estimation of
rainfall, if this technology would become available for measuring rainfall.

3.2.2 Some algorithms

One of the first empirical relations used was the relation between cloud top tempera-
tures (measured with a thermal band) and rainfall, also called 'cold cloud duration'. It
is now widely used with geostationary satellite imagery from Meteosat, Meteosat Second
Generation (MSG) and Geostationary Operational Environmental Satellite (GOES). For
instance the GOES Precipitation Index (GPI, Arkin & Meisner, 1987) is a precipitation
algorithm based on this relation. Cold cloud tops are caused by release of latent heat to
the atmosphere during convective rain storms and it is assumed that when a temperature
threshold is underspent, a certain rain rate will occur. Estimating daily rainfall is there-
fore a matter of counting the number of times that underspending of this temperature
occurred. This approach is therefore suitable for tropical areas, where most of the rainfall
is of a convective nature.

Microwave imagers (e.g. Special Sensor Microwave/Imager (SSM/I), Advanced Microwave
Soundings Unit (AMSU-B), TRMM Microwave Imager (TMI) and Advanced Microwave
Scanning Radiometer - EOS (AMSR-E)) may also be used to retrieve precipitation fields.
Both algorithms based on scattering (where the amount of scatter of the microwave signal
is dependent on the quantity and size of ice particles) and emission (where the variations
in brightness temperature are an indication of present water vapour and rainfall) have
been developed (e.g. Ferraro & Marks, 1995; Zhao & Weng, 2002). Problems with the first
type of algorithms occur where the rainfall is relatively 'warm' such as oceanic rainfall
events due to orographic lifting or shallow convective storms, because there are no or only
little ice particles present. For both algorithms, there are some problems with scale. The
footprint of a SSM/I sensor is in general too large to detect small convective events, since
the algorithms are based on footprint-averaged exceeding of thresholds.

TRMM is the first mission that carries a Precipitation Radar (PR) to estimate 3-dimensional
cloud properties and concurrent rainfall. More information on this satellite mission is given
in Section 3.2.3.

3.2.3 Merged rainfall estimates

The difference between rainfall estimates from some different algorithms is given in Fig.
3.3. It becomes clear that microwave algorithms see different rainfall properties than
cold cloud duration. Therefore, the best results are expected when different estimates are
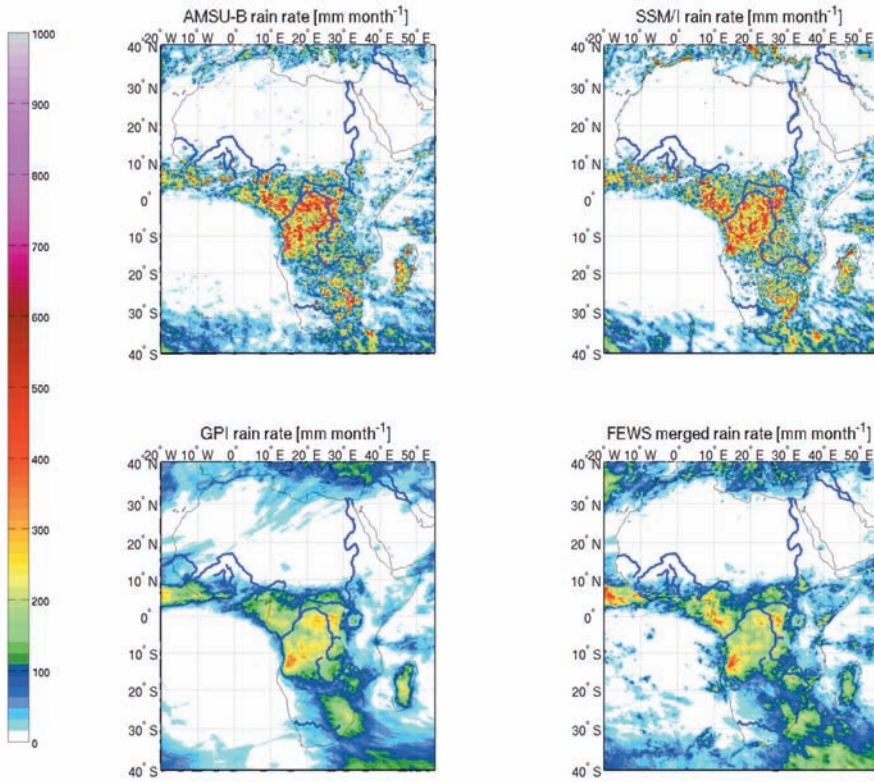
Figure 3.3: Three independent accumulated rainfall estimates and station data based estimates for a 30-day time window (November 1st until 30th 2007) over Africa. From left to right: AMSU-B microwave, SSM/I microwave, Meteosat GPI, and a merged estimate. The merged estimate shows the impact of rain gauges, especially over the Atlantic Ocean.

merged. In general, end user rainfall products are combinations of independent estimates, where the weight of the independent estimate is usually somehow based on local 'ground-truth' rainfall. Below, two popular rainfall estimates, also used in this thesis, are described.

FEWS RFE 2.0

A quite popular rainfall algorithm that is applied on a daily basis is the Rainfall Estimation (RFE) of the Famine Early Warning System (FEWS) version 2 (FEWS RFE 2.0). It is operationally and near-real time applied over several parts of the world, amongst which Africa. The strength lies in the fact that three independent rainfall estimations (due to the fact that three different instruments and relations are used) are merged together.

Each product is weighted inversely proportional to its variance with interpolated observed rainfall. This means that at locations where a lot of daily ground truth is available, the weight estimates are most meaningful, given that the ground-truth is of good quality. The weighting becomes artificial where the interpolation is more like an extrapolation due to the fact that the nearest observations are very far away. These spurious weights may result in random error, but also, if a remote ungauged area is subject to different rainfall mechanisms, in systematic errors.

Tropical Rainfall Measuring Mission (TRMM)

In 1997, the first satellite mission dedicated to observing tropical rainfall was launched. The Tropical Rainfall Measuring Mission (TRMM), carrying on board a microwave imager, infrared scanner and precipitation radar, has been flying for about 10 years now and will fly at least until 2009. Although this satellite mission is mainly designed for rainfall, its derived precipitation estimates are not independent of other satellites and ground data. The end products (e.g. 3B42 merged rainfall product) rely on a combination of the microwave, infrared and radar information from TRMM, other microwave sensors (AMSU-B, SSM/I and AMSR-E), infrared data from geostationary satellites and ground data, merged in the Global Precipitation Climatology Project (GPCP Adler *et al.*, 2003). Again, all estimates have their limitations and the best can be expected from a combination of estimates, anchored with ground-truth information (if such a thing exists), rather than one estimate alone.

Fig. 3.4 shows a comparison of the two described rainfall estimates over a number of years. Note that this comparison does not reveal which rainfall product is the best, it merely reveals differences and agreements between the two. The accumulated rainfall amounts (Fig. 3.4, top-panel) may be significantly different for both products. In some years FEWS seems to predict more rainfall and in other years TRMM. On annual basis, difference may be up to 10%, which is considerable, especially when one is interested in discharge. It is not clear what the cause of this inconsistency may be, but a hydrologist should at least be aware of this behaviour when using these data for modelling. An estimation of the spectrum (Fig. 3.4, bottom-panel) however, reveals a remarkable correlation between both products. Apparently, temporal and correlation signatures in the rainfall are very much alike considering both products, averaged over a basin. In the remainder of this thesis, these characteristics will be taken into account while considering how to profit optimally from remotely sensed rainfall estimates.
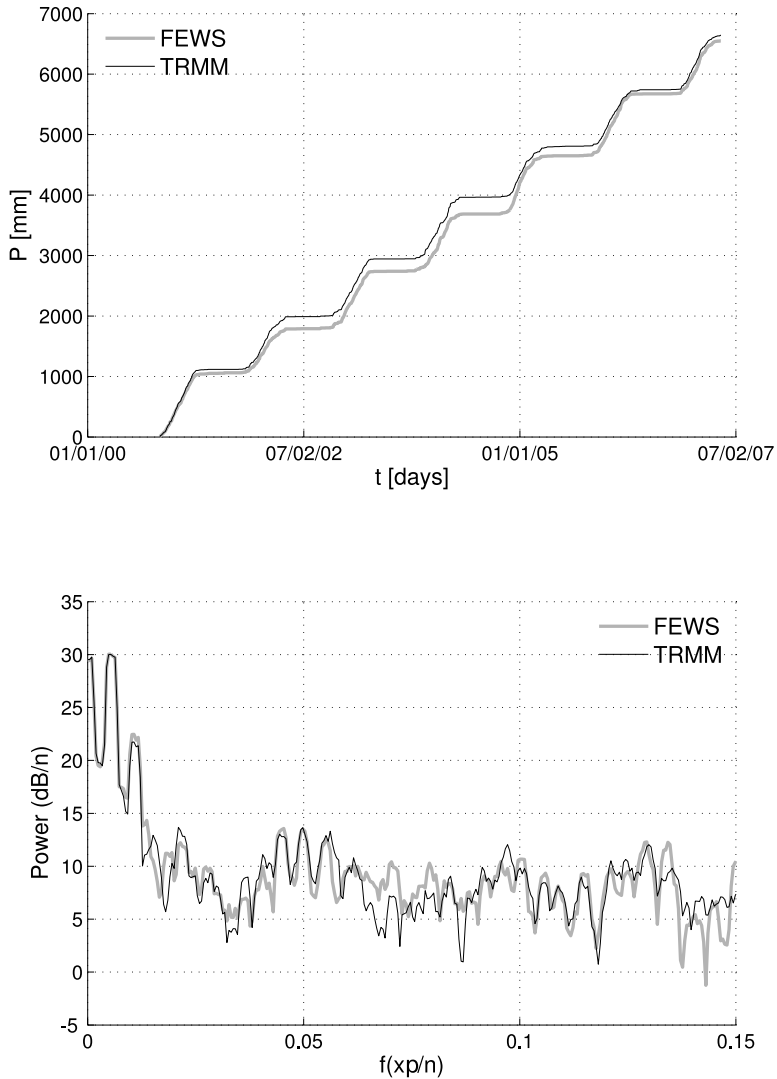
Figure 3.4: Top: accumulated rainfall from FEWS and TRMM over the period 2000-2007, averaged over the Luangwa river basin in Zambia (described in a later chapter). Bottom: periodogram of FEWS and TRMM data

3.3   Radiation from geostationary satellites

Evaporation is a key variable in the water balance of semi-arid river basins, often accounting for about 90% of the partitioning of rainfall at the land surface over long periods. Estimation of its temporal and spatial variability is therefore valuable in these areas, for instance for estimating the water balance where information on other fluxes is not

available (Bastiaanssen & Chandrapala, 2003) or for providing calibration data for hydrological models (Immerzeel & Droogers, 2008). Several algorithms have been developed to estimate the surface energy balance from satellite information, e.g. the Surface Energy Balance Algorithm for Land (SEBAL, Bastiaanssen *et al.*, 1998a; Allen *et al.*, 2007) and the Surface Energy Balance System (SEBS, Jia *et al.*, 2003). One of the major shortcomings of such algorithms is that during cloud overcast periods, no thermal information on the land surface can be retrieved from satellite images. To overcome this problem, often a Jarvis-type model (Jarvis, 1976) is applied using the governing meteorological conditions. Over large areas, the temporal variation of cloudiness and the related radiation budget at the surface is one of the key meteorological variables needed. Lack of information on the net radiation R_n may result in systematic error in evaporation estimates and thus the actual evaporation (E). Especially over Africa, it is necessary to infer evaporation under clouded conditions because of the prolonged rainy seasons, during which evaporation is mostly energy constrained (e.g. Schüttemeyer, 2005; Bagayoko, 2006). It means that if we wish to use evaporation as a reliable modelling constraint, both for spatial-temporal variability and as a water balance term, we are forced to also infer evaporation during cloud-overcast situations. Priestley and Taylor (1972) mentioned that knowledge of R_n and of ground-dryness constitutes a necessary prerequisite for the estimation of the main spatial variability in daily latent heat flux (λE) and sensible heat flux (H) on large scales. Makkink (1957) even proposed a simpler framework in which solely incoming short-wave radiation ($R_{s,in}$) is used to infer λE. Yet, little ground measurements of net radiation are performed in the world. Where measurements are performed, their quality is often compromised due to for instance dust particles that obscure pyranometers and lack of data paper for sunlight hour measurements in remote areas.

An obvious solution is to try to estimate incoming energy fluxes with satellite imagery. Of specific interest is incoming shortwave (or solar) radiation ($R_{s,in}$), because it is the major term of the net radiation R_n and the most variable in time. Satellites have the capability to estimate solar radiation over large regions through inverse modelling of the transmissivity, based on the reflectance properties of clouds in combination with the earth-sun orbiting geometry. Polar orbiting satellites have a limited overpass frequency, which is not enough to cover the temporal variability in transmissivity over the day. Fortunately, geostationary satellites have a high recording frequency and nowadays have a high spatial resolution. Combined with model-based estimates of cloud thickness, a reasonable estimate of incoming shortwave radiation can be made. In 2004, Meteosat Second Generation (MSG) was launched, to orbit the earth in a geostationary orbit at 0° latitude, 0° longitude, exactly covering Africa. Since then, the Land Surface Analysis Satellite Application Facility (LSA SAF) constructed an operational incoming shortwave

radiation product that can be downloaded from the internet (`http://landsaf.meteo.pt/`) and which has been used in this thesis. $R_{s,in}$ is largely dependent on the solar zenith angle and the atmospheric transmissivity. The cloud product of the Satellite Application Facility on support to Nowcasting and Very Short-Range Forecasting (NWC SAF) is used for cloud masking and separate algorithms for clear-sky and cloudy pixels have been adopted to convert the radiation at the top of the atmosphere to $R_{s,in}$. LSA SAF (2007) performed validation on several sites in Spain and Portugal which showed good performance under variable conditions.

The conditions in Sub-Saharan Africa are slightly different from the conditions in Spain. A unique condition is dust storms and significant rain storms. To ensure that these data are suitable under Sub-Saharan conditions, additional validation was carried out with data from some well-gauged sites. An extensive gauging campaign has been run for the GLOWA-Volta project (Van de Giesen *et al.*, 2002) in the Volta river basin (see `http://www.glowa-volta.de/`). This basin has a pronounced gradient in climate and land cover. Within the basin area, rainfall ranges from 700 mm yr^{-1} in the North to 1500 mm yr^{-1} in the South. Two of the project meteo-stations (see Table 3.1 and Figure 3.5) were used to evaluate data sources derived from MSG. The meteo-stations cover a varying land cover and climatology.

Table 3.1: Location of meteo-stations in the Volta basin.

Meteo stations	Latitude	Longitude	Altitude [m]
Dano	11° 09' 42.20" N	03° 04' 34.10" W	331
Ejura	07° 20' 29.40" N	01° 16' 38.75" W	208

Dano

Dano is located in the South-Western part of the Volta Basin. Terrain is generally mildly sloped and covered by laterite plateaus (270 - 350 m above sea level). Landscape and vegetation is classified as Sudan savannah belonging to the Sudanian-Guinean domain. The test site vegetation is comprised of maize, peanuts, sorghum, and millet. Mean annual rainfall according to ARTES climatology (1961 to 1990) is ~930 mm yr^{-1}. Rainfall follows a mono-modal pattern with a peak in August. The wet season lasts for roughly 4 to 6 months with a core time from June to September. The onset of the rainy season underlies high variability.
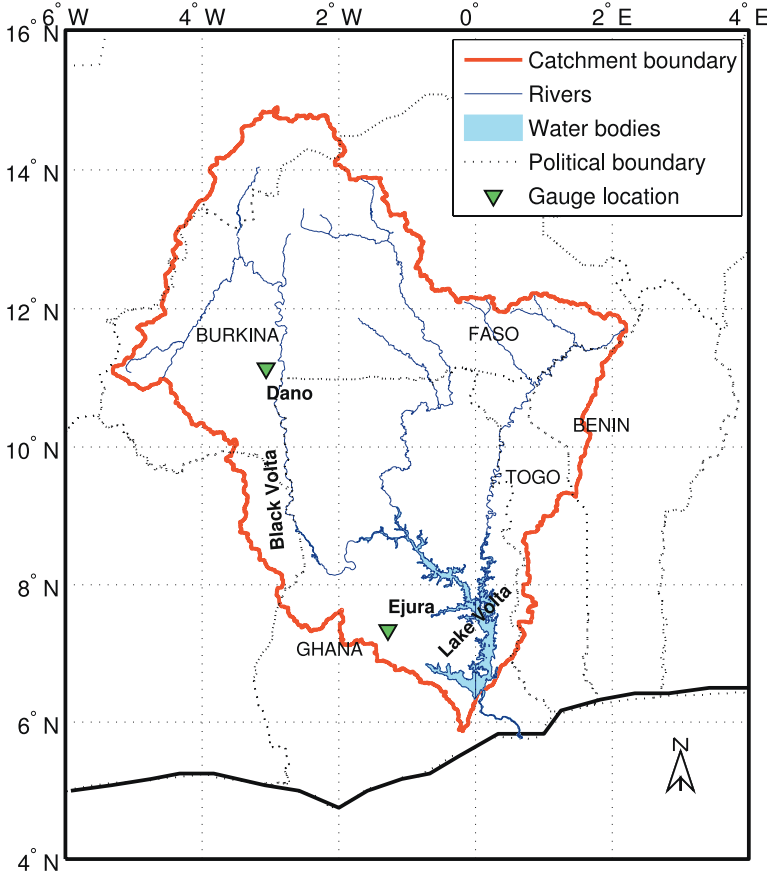
Figure 3.5: The meteo stations used, Dano and Ejura, located in Ghana and Burkina Faso

Ejura

Ejura is situated on the southern fringes of the Volta Basin in a slightly hilly terrain (150 - 250 m above sea level). It lies in the transition zone between forest savanna and mountainous guinea forest. The test site is located around a valley with vegetation covering local food crops, such as yams, corn, cassava; shrub, including elephant grass; and trees, amongst them cashew, various palm trees, and banana trees. The valley bottom is swampy and the elevation differences are around 100 m. Mean annual rainfall is about 1400 mm yr^{-1}. Rainfall follows a pseudo bimodal pattern, meaning that in August it experiences a slight decrease, and the peak of the season is in September. The wet season lasts for roughly 8 months from March till October, whereas the onset of the season underlies high variability. This site is subject to far more cloud cover than Dano.

Validation was carried out by extracting a time series of MSG-derived $R_{s,in}$ for pixels,

wherein the two meteorological stations are located. These 30-minute estimates were compared with 10-minute in-situ records having the same time stamp as the MSG estimates. $R_{s,in}$ can be validated on this scale, because the variability of transmissivity is largely dependent on variability in cloud coverage. This is a large scale process and will not be significantly variable within the scale of an MSG pixel ($\sim 3 \times 3$ (km)^2).
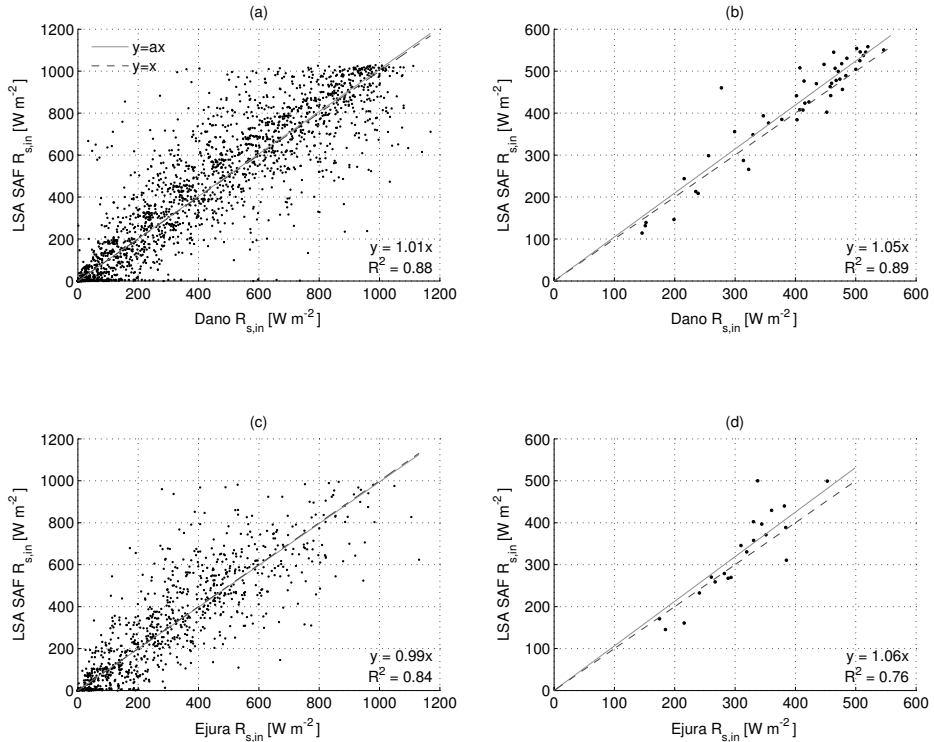
Figure 3.6: Left: Half-hourly scatter between in-situ measurements in Dano (a) from 8 July 2005 until 12 October 2005 and Ejura (c) from 8 July 2005 until 21 August 2005. The right side presents daily-averages in Dano (b) and Ejura (d). The selected rather short time windows are chosen because LSA SAF started disseminating $R_{s,in}$ estimates on 8 July 2005 and the available in-situ records do not go further than 12 October 2005 and 21 August 2005 for Dano and Ejura respectively.

Fig. 3.6 shows a scatter plot of the resulting comparison. The results show that there is almost no bias between in-situ records and LSA SAF estimates. At Dano, the explained variance is 88% on half-hourly basis and 89% on daily averages. In Ejura, explained variance is 84% on half hourly basis and 76% on daily basis. The LSA SAF estimates seem to saturate at about 1000 W m^{-2}, which may indicate that they are not suitable when radiation exceeds this value. This is not likely to occur frequently except in the Sahara desert.

3.4   Data from atmospheric models

Atmospheric models are capable of generating global estimates of the atmospheric states and fluxes based on initial states and analysed data, for instance rawinsonde measurements. An example of a long time series, generated by assimilation of satellite and in-situ (rawinsonde) observations and an atmospheric model, is the European Centre for Medium range Weather Forecasts (ECMWF) 40-year Reanalysis project (ERA40, Uppala *et al.*, 2005), which provides a consistent dataset of the state of the atmosphere and land surface from 1957 until 2002. Data fields from this reanalysis and the operational ECMWF model have been used in this thesis as input to satellite based surface energy balance estimates. To this end, downscaling of the fields, which are typically $\sim 100 \times 100$ (km)^2 is necessary. The next section describes the downscaling methodology, developed during an MSc study and enhanced during this thesis.

3.4.1   METEOLOOK: a physically based downscaling approach for meteorological measurements

Evaluation of the energy balance with satellite images requires high resolution grids of wind speed, air temperature and humidity. In remote areas these data are often not readily available. While the spatial variability of relative humidity may be assumed to be small, the distribution of temperature and wind speed are greatly influenced by the land surface. Voogt (2006) developed a method which allows for physically based distribution of sparse in-situ measurements, called METEOLOOK. The method is based on the Monin-Obukhov theorem (see e.g. Brutsaert, 1982). Based on sparse in-situ measurements, the roughness properties and a first approximation of the energy balance (based on Normalised Difference Vegetation Index, NDVI) at these sites, vertical wind and temperature profiles are estimated. Hence, temperature and wind can be computed at 'blending' height above the measurement point, arbitrarily fixed at 100 m above the land surface. Here, it is assumed that temperature and wind speed are no longer influenced by the land surface directly beneath it, but only by large scale features such as elevation, distance to sea and the regional radiation budget. Variability of temperature and wind due to these features is estimated and corrected for. The residuals are geostatistically interpolated and transformed to near-surface variables using the inverse of the previously described steps.

In this study, in-situ measurements were not readily available everywhere, more specifically in the case study described in Chapter 7. In-situ measurements, concomitant with the required MSG radiation information were not available. Therefore, Voogt's approach was extended iKeywords: n this PhD study to be applicable to atmospheric modelled data fields and has been applied on ECWMF model fields. This work has not been published

and is therefore elaborately described below. More information can be found in Voogt (2006).

3.4.2 METEOLOOK applied on atmospheric modelled data

An atmospheric model field level is selected that approximately matches a subjective 'blending' level of approximately 100 m. At this height above surface, the temperature and wind speed are assumed to be 'blended', meaning that the land surface does not influence its spatial variability a lot. In the case of the operational ECMWF model (60 full levels, extended to 92 in February 2006) this level would be terrain following model level 58 (before Feb. 2006) or 90 (after Feb. 2006), which over Africa generally follow the terrain at an elevation of about 92 m. In the remainder of this section, the subscript 'm' is reserved for model derived variables at this model level. Model variables u_{e-w} (east-west wind speed), u_{n-s} (north-south wind speed), p_m (air pressure), q_m (air specific humidity) and T_m [K] (air temperature) from this level are needed. u_{e-w} and u_{n-s} are combined with Pythagoras' law to yield u_m, the effective wind speed [L T^{-1}] over a ECMWF model grid cell. Pressure at model level follows from

$$p_m = a + bp_{s,m}, \tag{3.1}$$

where a and b are fixed parameters [-] and $p_{s,m}$ [M L^{-1} T^{-2}] is the pressure at the land surface, also retrieved from ECMWF. The model level height then follows from

$$z_m = \frac{\varrho T_m}{g} \ln \left[\frac{p_m}{p_{msl}} \right], \tag{3.2}$$

where ϱ [M^2 L^2 T^{-2} K^{-1}] is the universal gas constant, p_{msl} is an assumed constant pressure at mean sea level, g is the gravitational acceleration [L T^{-2}], z_m [L] is the height of the model level above sea level. Now that all model information is available, the distribution can be performed in the following 3 steps:

1.   The spatial distribution of temperature (both instantaneous and period-averaged) at blending level is assumed to be linearly related to elevation and incoming solar radiation within the region of interest. With linear regression, this relation is established for each field considered, with (ECMWF) pixel-effective elevation and pixel-averaged solar radiation (derived from MSG). The relation is then applied on a full resolution 1×1 (km)^2 Digital Elevation Model (DEM) and MSG radiation, using the ECMWF blending level temperature field as 'measurement', to create a grid that represents the temperature at a constant height (blending level z_b) above the true surface level, instead of above the artificial, effective model surface level (z_m), represented by the geopotential height of the ECMWF model:

$$T_b(z_b) = T_m(z_m) + \beta_1 \tau R_{s,ex} + \beta_2 \Delta z, \tag{3.3}$$

where $T_m(z_m)$ and $T_b(z_b)$ [K] are the original ECMWF temperature and 1×1 (km)^2 distributed temperature at blending level, τ [-] is the transmissivity of the atmosphere (derived from MSG), $R_{s,ex}$ is the extra-terrestrial solar radiation [M T^{-3}], Δz [L] is the elevation difference between the effective ECMWF elevation and a high-resolution DEM, and β_1 and β_2 are regression constants. Note that the latter is in fact a lapse rate, also used for downscaling in Cosgrove *et al.* (2003).

2.   The computation of near-surface temperature and wind speed (where the land surface is of influence) is established through the use of Monin-Obukhov's theory (see e.g. Brutsaert, 1982). This theory relates the vertical profiles of wind speed and air temperature with the sensible heat flux and the surface roughness through relations

$$u_\star = \frac{\kappa u_m(z_b)}{\ln(z_m/z_{0m}) - \psi_m(z_b/\Lambda) + \psi_m(z_{0m}/\Lambda)}, \qquad (3.4)$$

with

$$\Lambda = -\frac{\rho_a c_p T u_\star^3}{\kappa g H}, \qquad (3.5)$$

where Λ [L] is the Obukhov length for atmospheric stability, $u_\star$ is the friction velocity [L T^{-1}], κ [-] is the von Karman constant, z_{0m} [L] is the momentum roughness height, ψ_m is an empirical stability correction function for momentum, ρ_a is air density [M L^{-3}] and c_p is specific heat of air [L^2 T^{-2} K^{-1}]. For the computation of ψ_m, the reader is referred to Brutsaert (1982). Unknowns are H, Λ and $u_\star$. If H can be approximated, Λ and $u_\star$ can be determined iteratively through Eqs. 3.4 and 3.5. H is estimated through a first proxy of the energy balance. An empirical relation between the energy balance and the Normalised Difference Vegetation Index (NDVI) and the vapour pressure deficit is assumed. For details, see Voogt (2006).

3.   In the final step, eq. 3.4 is inverted to compute wind speed at reference (near surface) level z_{ref}:

$$u(z_{ref}) = \frac{u_\star}{\kappa} \left[\ln\left(\frac{z_{ref}}{z_{0m}}\right) - \psi_m(z_{ref}/\Lambda) + \psi_m(z_{0m}/\Lambda) \right].$$

Finally temperature at reference level, $T(z_{ref})$ is computed by

$$T(z_{ref}) = T(z_b) - \frac{T_\star}{\kappa} \left[\ln\left(\frac{z_b}{z_{ref}}\right) - \psi_h(z_b/\Lambda) + \psi_h(z_{ref}/\Lambda) \right], \qquad (3.6)$$

with

$$T_\star = -\frac{H}{\rho_a c_p u_\star}.$$

Here ψ_h is a stability correction function for heat and is computed following Brutsaert (1982).

3.4.3 Validation
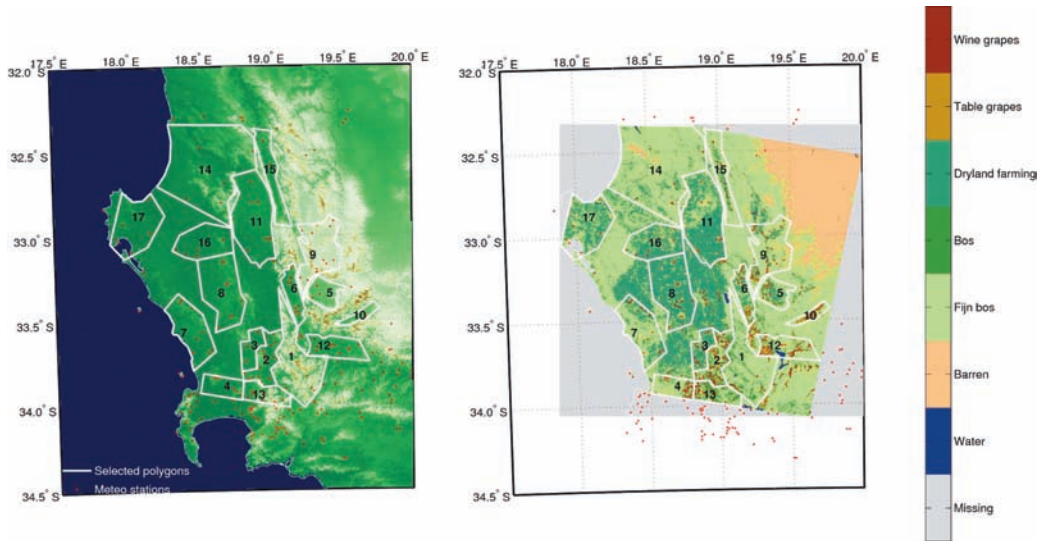
Figure 3.7: Western Cape validation area. The polygon shaped areas are delineated on the basis of similar land cover and elevation. The left figure shows the delineated polygons overlaid on an elevation map, the right figure shows the same, but overlaid on a LandSAT classification of land cover (source: WaterWatch, Wageningen).

Validation of the model-based version of METEOLOOK has been carried out over a relatively well monitored area around the South-African Cape (Fig. 3.7). Ground station data has been obtained by WaterWatch associates from the Agricultural Research Council (`http://arc-agric.za`) and the South-African Weather Service (`http://www.weathersa.co.za`). The set of stations consists of a mixture of automatic recording stations and synoptic manual stations. Twenty-four hour averaged data have been used to evaluate the effectiveness of METEOLOOK in downscaling both temperature and wind speed. This has been done over one month of data (September 2005) over delineated polygon-shaped hydrotopes, shown in Fig. 3.7. These hydrotopes were selected on the basis of similar land cover and elevation. Unfortunately, the station density, shown in Fig. 3.7 is time variable, which means that in general the station coverage is much poorer than depicted in this figure. For each day, measured values were averaged per polygon and compared with the pixel-averaged values of the near-surface ECMWF temperature and wind speed, as well as the high-resolution METEOLOOK downscaled temperature and wind fields.

In ungauged areas, surface roughness may be related to land use maps. To simulate the ungauged case, this was also done over the Western Cape. The International Geosphere

Biosphere Programme land cover map (IGBP, 1990) was used to assign roughness values for each land cover type.

Table 3.2 shows some statistics on the validation. To emphasise the meaningfullness of the validation, the station density during September 2005 is also displayed in the second column. Downscaled wind speed does not compare better with measured wind speed. This may partly be explained by the simplified representation of roughness. The large problem with roughness over agricultural areas, is that it is strongly seasonal, which may lead to spurious results. Another problem is the spatial variability of wind speed, which may be influenced by sheltering which is often even dependent on the wind direction. Wind speed therefore remains a variable, which is difficult to validate. Temperature estimates, however, are greatly improved by METEOLOOK. Both root mean square error and bias reduce significantly by application of the algorithm in almost all polygons. In some polygons, there is a low resemblance between measurements and METEOLOOK, which can be explained by the low availability of ground stations in September 2005, which reduces the area-representativeness of the measurements.

Table 3.2: Performance comparison of ECMWF at the lowest model layer (z_1) without downscaling ($u(z_1)$ and $T(z_1)$) and with downscaling ($u(z_{ref})$ and $T(z_{ref})$) for wind speed and temperature near the land surface. Performance indicators root mean square error (RMSE) and bias (B) were used.

Wind speed [m s^{-1}]

Pol.- elev./land cover	# stat.	RMSE $u(z_1)$	RMSE $u(z_{ref})$	B $u(z_1)$	B $u(z_{ref})$
1 - high/fyn bos	0	-	-	-	-
2 - low/wine	15	0.88	1.15	0.81	-0.94
3 - low/dryland	3	1.39	0.82	1.28	-0.50
4 - coast/fyn bos	5	2.00	1.21	1.91	-0.81
5 - low/dryland	0	-	-	-	-
6 - low/dryland	3	0.49	0.53	0.44	-0.37
7 - coast/fyn bos	3	3.62	2.05	-3.19	-1.76
8 - low/dryland	3	0.90	1.63	0.71	-1.43
9 - high/fyn bos	1	0.79	1.40	0.73	-1.24
10 - low/wine,table	3	1.30	0.65	1.17	0.19
11 - low/dryland	2	0.77	0.53	0.64	-0.20
12 - low/fyn bos	3	0.50	1.15	0.28	-1.09
13 - mixed/mixed	5	1.50	0.64	1.34	0.39
14 - coast/fyn bos	3	0.91	1.59	0.77	-1.35
15 - low/dryland,wine	1	0.78	0.88	0.69	-0.66
16 - low/dryland	2	3.44	2.19	-3.18	-2.00
17 - coast/dryland,fyn bos	1	2.25	1.84	-1.83	-1.34

Temperature [°K]

	# stat.	RMSE $T(z_1)$	RMSE $T(z_{ref})$	B $T(z_1)$	B $T(z_{ref})$
1	0	-	-	-	-
2	15	2.54	1.52	2.37	0.83
3	3	2.31	1.23	2.17	-0.08
4	5	1.93	1.44	1.74	0.93
5	0	-	-	-	-
6	3	2.31	2.22	2.02	-1.24
7	3	1.87	1.45	0.19	-0.43
8	3	2.03	1.40	1.43	-0.42
9	1	1.71	3.65	-0.67	-3.33
10	3	2.26	1.19	-2.00	-0.48
11	2	3.43	2.23	2.66	0.19
12	3	1.82	3.16	0.13	-2.28
13	5	3.09	2.11	2.64	-0.36
14	3	2.17	1.42	1.92	0.63
15	1	1.64	3.25	0.52	-2.59
16	2	1.58	2.00	-0.45	-1.66
17	1	2.09	1.45	1.18	0.52

3.5 Synthesis

This chapter has dealt with some of the information sources, that are used throughout this thesis. Most of them are available to practically any user that is able to handle them. They are all free of charge, except for the information of the atmospheric model, ECMWF. The dissemination of these data is generally limited to meteorological institutes and institutes and companies under conditions of a memorandum of understanding or a contract with a national meteorological institute. It is especially interesting to realise that during the start of this thesis work, some of the mentioned data sources were not available yet. The Meteosat derivatives for land surface applications, mentioned in section 3.3, became limitedly available in July 2005, after which the availability increased to almost continuous. Furthermore the list of information provided here could be extended with near-real time available Normalised Difference Vegetation Index (NDVI), soil moisture, lake and river altimetry, snow cover and so on. These data were not explicitly used for model improvement in this thesis and were therefore omitted from this overview. The chapter showed that all satellite-based estimates of variables, related to hydrology, may be subject to significant uncertainty, and possibly bias. In the case of rainfall, radiation and, described in a later chapter, evaporation, this has partly to do with the fact that satellite instruments are not capable of directly measuring fluxes. Instead, a state is measured (i.e. the cloudiness, cloud temperatures, land surface temperature, droplet properties etc.), which is consequently transferred into the variable of interest by means of a transfer model. GRACE is perhaps the only satellite mission that observes a phenomenon that is directly equivalent to a variable, generally used in the water balance of a hydrological model, but here the limited spatial and temporal resolution plays a major role in its applicability, and its accuracy is undeniably influenced by instrument noise and the quality of background models.

Summarising, every remotely sensed data source contains information, but the issue here is to learn how to extract the information in the data by means of modelling, while taking into account their uncertainties. These uncertainties are often difficult to quantify, which leads to the conclusion that remotely sensed estimates of hydrological variables should be regarded as soft data. Throughout this thesis, the softness of remotely sensed data is considered as much as possible. Below, an overview is given of the data sources mentioned in this chapter, along with their temporal and spatial resolution, possible uncertainties and how they may be used in hydrology.

Table 3.3: Summary of available remote sensing data, spatial and temporal scale, sources of uncertainty and availability

Data source	Spatial scale (km)^2	Temporal scale	Error sources	Use
GRACE (water storage variation)	$\sim 200 \times 200$	$\sim$ Monthly*	Measurement noise, background models, leakage, resulting in scale dependent uncertainty	Validation
Satellite rainfall	$\sim 25 \times 25$	$\sim$ (Sub)daily	Aliasing, model errors, weights, resulting in bias and noise	Input water balance models
Radiation	$\sim 3 \times 3$	~ 30 minute	Model errors, measurement noise	Input energy balance models
Downscaled meteo	$\sim 1 \times 1$	Instantly or period averaged	Model errors	Input energy balance models

---

*. GFZ, NASA and CNES are also producing weekly solutions but at the cost of a considerably lower spatial resolution, in the order of 1000 km.

Chapter 4

Analysis of monthly GRACE data over Southern Africa

In order to become a useful data source for hydrological modelling at the basin scale, error sources of GRACE that may have a significant effect on its information content for hydrology should be considered. In this chapter, a short overview of these error sources is given. Important errors are introduced in the monthly GRACE solutions due to *aliasing*. This phenomenon is caused by the fact that the sparse sampling capacity of GRACE picks up high frequency processes, which can only partly be compensated for, by subtracting model predictions of these processes from the GRACE signal. As a result, a *spatially correlated error structure* is revealed in the monthly solutions. The sparse overpass behaviour may cause aliasing in the time stamp of the monthly solutions in specific river basins as well. Furthermore, spatial filtering, needed to reduce the aliasing errors, results in *leakage errors*. The significance of these error sources is demonstrated over the upper Zambezi. Concluding remarks from this chapter have been used during further study. Henceforth, in the next chapter it is shown that hydrological models, even if they are not fully calibrated, can be used to improve GRACE basin-scale solutions, which in turn may be employed as a complementary means of validation of hydrological models. This is demonstrated in the upper Zambezi and surrounding river basins in the next chapter.

4.1 Error sources of GRACE and their effect on hydrological usefulness

4.1.1 Aliasing

The GRACE twin satellites do not have a fixed orbit as for instance many polar orbiting satellites with visual or thermal measuring instruments have. In fact, the whole idea is that the satellites move freely, having an orbit purely determined by gravitational forces. Therefore, GRACE orbits (or overpass tracks) are irregularly spaced in time and have a limited overpass frequency. Furthermore, GRACE measurements are sensitive mainly to variations along the ground-tracks, which limits the sampling of east-west oriented mass redistributions. The signal that GRACE is sampling consists of many low frequency, but also high frequency processes. The high frequency processes include tides and oceanic and atmospheric mass redistribution. Their effect on the gravity solutions is reduced by subtracting model predictions of these phenomena from the solutions. Imperfect as these models usually are, some high frequencies will still contaminate the solutions. This is an aliasing problem and causes the along-track oriented 'striping' in the gravity solutions (see e.g. Fig. 4.1 top-panel). Han *et al.* (2004) quantified the effects of aliasing due to background model error and concluded that some of the aliasing errors could be up to 3 times the expected measurement noise in the monthly solutions. Seo *et al.* (2008) investigated which coefficients are likely to be affected by aliasing and concluded that errors in estimates of temporally high frequency processes could cause aliasing, affecting low frequency coefficients of the spherical harmonic coefficients.

Another effect of aliasing has been observed by Winsemius, Savenije, Van de Giesen, *et al.* (2006). They hypothesize that the irregular spacing of the satellite tracks may also regionally result in aliasing of water storage estimates, when the information collected during the orbits is equally weighed and averaged in both space and time. The extreme case may occur during periods when the GRACE satellites are in an almost repeat orbit (e.g. in fall 2004). When this occurs, the inter-track spacing is relatively large and remains so during a long period. hence a target area is poorly sampled, which was also later shown by Klokočnik *et al.* (2008). This may cause loss of information about the temporal variability of both the true storage variation and the orbit behaviour over, and in the neighbourhood of specific target areas. Naturally, such problems gain importance when the selected target area becomes smaller.

4.1.2 Correlated errors

The previous section described that aliasing errors manifest themselves as North-South oriented stripes. The consistency of these stripes suggests that the aliasing errors are

spatially correlated. This has been investigated in detail by Swenson and Wahr (2006). They developed a post-processing filter, which smooths correlated Stokes coefficients in the spectral domain as to reduce the effects of correlated errors. The optimal filter, described by Klees *et al.* (2008), also takes care of correlated features such as the North-South oriented stripes by using a non-symmetric and anisotropic filter, which is variable for each monthly solution, dependent on the full variance-covariance matrix of the individual monthly solution.

4.1.3 Leakage, due to spatial filtering

As described in the previous sections, GRACE estimates of monthly mean water storage variations are subject to errors due to measurement noise, propagation of errors in the used background models, and consequently the aliasing of imperfectly estimated high-frequency mass variations into the monthly GRACE gravity field solutions (Swenson & Wahr, 2002; Wahr *et al.*, 1998, 2006). Measurement noise and errors in background models, used to remove effects of oceanic movements, tides and atmosphere causes 'stripes' in the GRACE monthly solutions (i.e. the individual satellite overpasses are visible). In fact, at small spatial scale (i.e. at high spectral frequencies), noise may have a far greater variability than the signal. We can reduce these errors by spatial filtering (i.e. smoothing), which is in fact routinely applied. Fig. 4.1 shows as an example GRACE storage change between March 2003 and April 2003, both unfiltered (top) and filtered with a Gaussian filter, half-radius 700 km (bottom).* The colour scale of both sub-figures is different by a factor of 5 to emphasise the signal-to-noise ratio. Unfortunately, spatial filtering biases the GRACE estimates of monthly mean mass variations, because smoothing over a discrete spatial shape (in our study cases, a river basin) means that some information from outside the shape filters into the shape and some information from inside the shape filters out of the shape. This error is referred to as *leakage* in this thesis. It is the subject of optimal filter design to find a filter that minimises the sum of GRACE errors and filter error. In fact, a parallel study has been undertaken by DEOS to prepare such an optimal filter (Klees *et al.*, 2008). As a result, an anisotropic, time-variable filter has been prepared that optimises the north-south orientation of the twin-satellites (i.e. more filtering in north-south direction than in west-east direction) and uses the time-variable variance-covariance information of the monthly solutions (i.e. a surrogate for the total error due to noise and aliasing of high frequency mass variations) with excellent results. In practise, designing an optimal filter would require the exact knowledge of the signal and noise autocovariance functions.

―――――

*. The solutions were provided by the Department of Earth Observations and Space systems (DEOS), TU Delft.
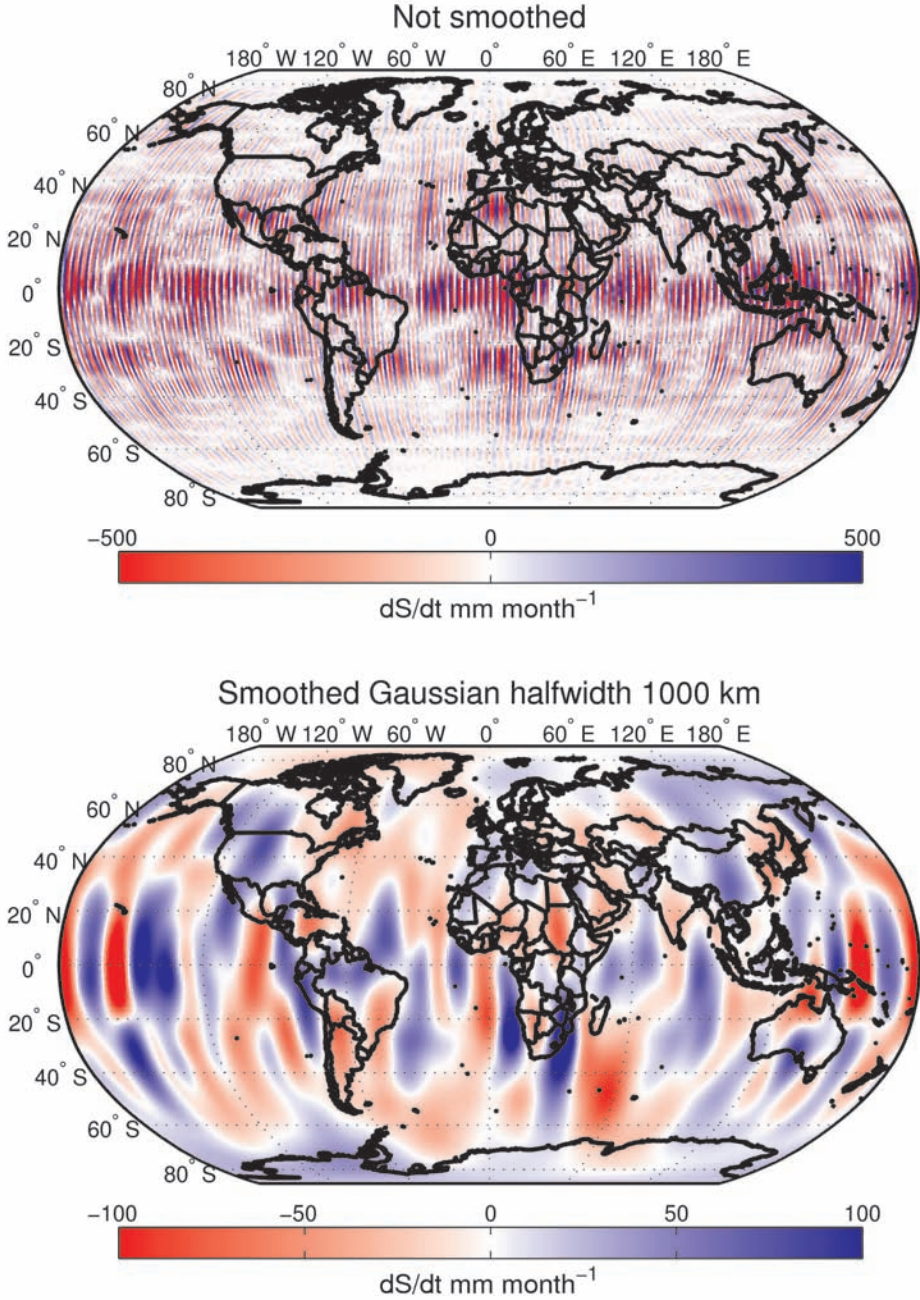
Figure 4.1: Global monthly mean water storage change between 17 March and 16 April 2003, solutions from DEOS, TU Delft. Top: non-smoothed. Bottom: smoothed with a Gaussian filter, correlation length, 700 km.

4.1.4 Effect of errors on hydrological signal

Various studies have compared terrestrial water storage data at river basin scale derived from GRACE with other datasets. Rodell, Famiglietti, *et al.* (2004) and Andersen *et al.* (2005) validated GRACE estimates of temporal variation in storage, in this chapter referred to as $\Delta S/\Delta t$, over the Mississippi and Europe, respectively. They compared them with estimates inferred from a combined atmospheric and terrestrial water balance using 40-yr ECMWF Re-Analysis (ERA-40) derived atmospheric moisture convergence (Seneviratne *et al.*, 2004), with soil moisture data from the Global Land Data Assimilation System (GLDAS, Rodell, Houser, *et al.*, 2004) and in-situ gravimeter measurements. The comparison between these datasets generally shows agreement. However the temporal pattern is often not entirely comparable, although this is not acknowledged by the authors. Inconsistencies between GRACE and the other datasets in timing of the up- and downgoing storage can often be noticed. A possible reason may be that land surface models, such as GLDAS, only represent soil moisture in the unsaturated zone, while storage variation in groundwater and wetlands, which has a significantly different timing and memory, is not considered. GRACE, however, gives the total profile of water storage change in the earth. Andersen *et al.* (2005) also left open the question whether the comparison of the coarse scale and spatially filtered solution of GRACE with ground observations at the point scale is valid.

Apart from these shortcomings, the aforementioned aliasing errors and resulting correlated errors and leakage should be considered when studying the usefulness and shortcomings of GRACE in hydrological modelling. Errors, due to leakage are for a large part a function of spatial variability of the storage variation in time. In the extreme cases, the spatial variability is represented by a sharp boundary at the edges of the catchment (e.g. where a catchment borders an ocean), or by no spatial variability at all. In the former case, the effect of leakage, relative to the storage, may be significant. This is especially the case when the catchment is small compared to the spatial scale of the filter, or when the storage signal in the basin of interest is out of phase with the storage signal of its surroundings. In the latter case it will be zero. Over a hydrological year, smoothing will obviously lead to a reduction in temporal variability of the storage in any target area. In view of the studies by Rodell, Famiglietti, *et al.* (2004) and Andersen *et al.* (2005), the temporal inconsistencies could be the result of spatial smoothing. This is because the leakage error for each target area is not constant in time. In fact its temporal variability is strongly dependent on the amount of spatial variability of storage in each month, relative to the mean gravity field. The aliasing due to irregularly spaced overpasses is also likely to result in temporal inconsistencies. Below, these effects are demonstrated, using monthly pre-processed GRACE gravity fields. For this preliminary analysis, gravity field

solutions from February 2003 until January 2006 were taken from GeoForschungsZentrum (GFZ) Potsdam's release 03. These data have been corrected for oceanic and atmospheric contributions and converted to equivalent water storage heights. Note that techniques for removal of correlated errors were not available at the time of study.

Effect of spatial filtering
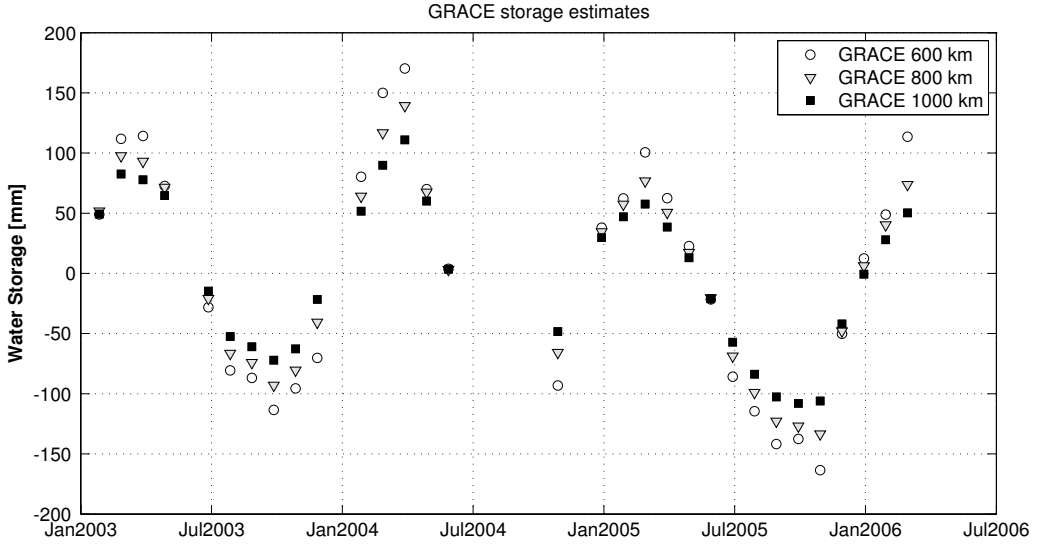
Figure 4.2: GRACE storage estimates of the Zambezi, upstream of Victoria Falls, after applying different Gaussian filters. The estimates are based on GFZ Release 03 GRACE models.

A routinely applied spatial filter has been applied to smooth GRACE results and retrieve storage estimates for the upper Zambezi river basin (its outline being given in Figs. 4.3 and 4.4). A Gaussian filter was used where the convolution matrix may be described by

$$f(r) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{4.1}$$

and a spatially filtered estimate of storage S_f can be obtained from the original unfiltered GRACE storage S_g by

$$S_f(r) = S_g(r) * f(r). \tag{4.2}$$

Gaussian-type filters with three different half radii (600, 800 and 1000 km) have been applied to the GRACE monthly solutions. The half radius r_h may be described as the distance from the center of the filter kernel to where the filter has lost half of its power

and may be described as

$$r_h = \sqrt{\left|\left[2\sigma \ln\left(\frac{1}{\pi\sigma^2 f(0)}\right)\right]\right|}. \tag{4.3}$$

Subsequently the area averaged storage estimates over our target area have been compared. Figure 4.2 shows the effect of the filters on GRACE storage estimates. It demonstrates that, besides dampening the annual variability of the signal, the filtering may slightly alter the apparent temporal signature of both datasets. This is because in some months the spatial variability in storage with respect to the surroundings of the target area may be larger than in other months. In this case, the filter causes more spatial attenuation than when this variability is small. This can be observed best in the vicinity of the storage extremes (i.e. lowest storage and highest storage with reference to a long term mean). The consequence is that with a large filter radius, not only the amplitude of annual storage variation is reduced, but also interannual variability is reduced, which reduces the information content of the data with respect to hydrological applications. 2004 for instance was clearly a wetter year than other years, for the upper Zambezi. This is only marginally apparent when using a 1000 km half radius. Furthermore, between February and March 2003, the GRACE time series shows that the the sign of storage change, changes from positive to negative when a larger filter radius is applied.

Rainfall and ground tracks

To illustrate how aliasing may influence GRACE monthly solutions, we have to study how GRACE really samples our target area within one month and how this relates to within-month hydrological processes. To this end, the spatial and temporal rainfall patterns within a month were investigated, in combination with ground track patterns. The FEWS RFE 2.0 rainfall maps mentioned in Chapter 3 have been used. Each observation period of approximately one month has been split up in 9 bins of about 3 or 4 days during which the ground tracks over the target area as well as the net rainfall (P_n) are jointly mapped. Net rainfall is defined as rainfall, exceeding a threshold of 2 mm day^{-1} (E_i) to account for interception, the dominating high frequency hydrological process, according to the following equation (Savenije, 1997, 2004).

$$P_n = \max(P - E_i, 0). \tag{4.4}$$

In this way, we can visualise what the effect may be of the spatial and temporal variability within a month of both the net rainfall and the GRACE orbit behaviour on the final monthly GRACE solution. Here, two disaggregated months have been visualised, where different filter correlation lengths clearly result in a different storage pattern, according to the results in Fig. 4.2. For display purposes, only the ground tracks within the vicinity

of about 500 km from the upper Zambezi are shown. In September 2003, Fig. 4.2 shows
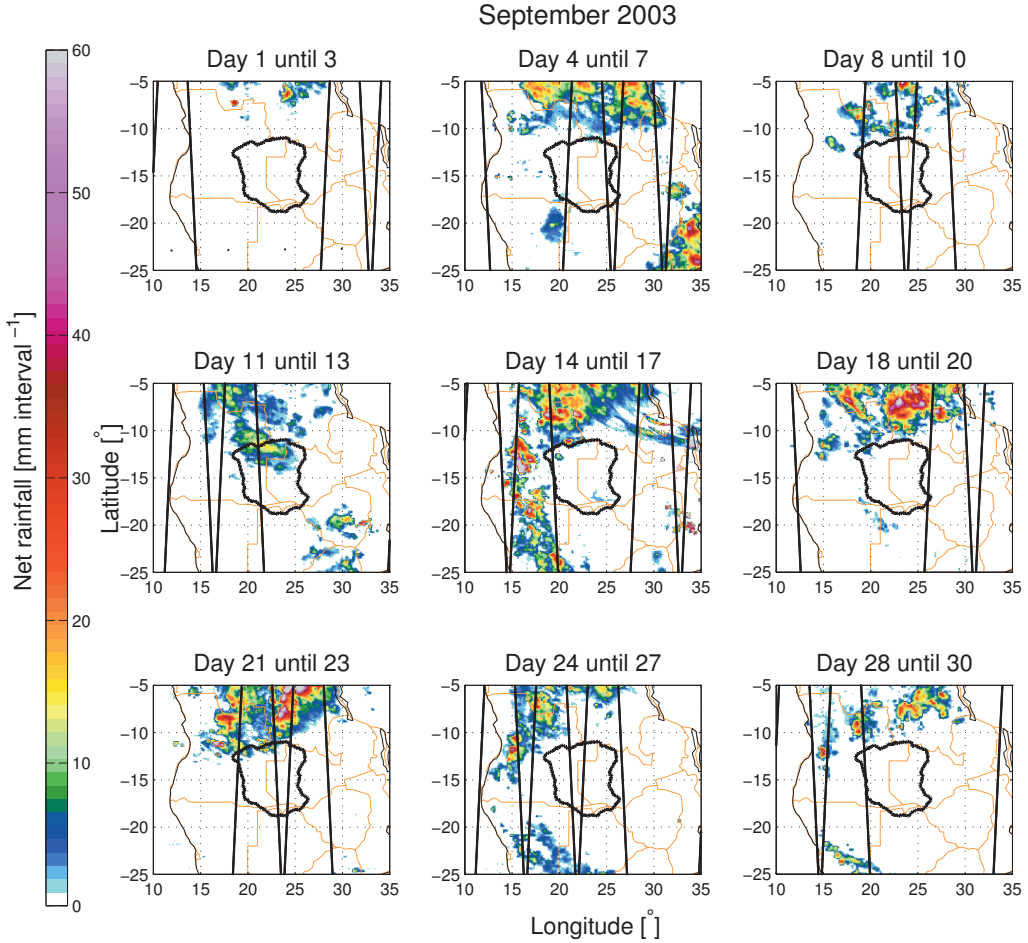
Figure 4.3: Ground track coverage during 9 intervals of 3-4 days in September 2003. The target area is delineated in black. Black dots indicate the ground tracks.

a slight increase in dS/dt when a relatively small filter correlation length is applied. This is an unexpected behaviour because the amount of rainfall in this month is negligible and storage change seems to decrease further in October 2003. The longer filter lengths do not show this increase. Fig. 4.3 shows the GRACE observation period in September 2003. From the 9 displayed bins, we can already observe how irregularly spaced the ground tracks are. Evidently, September is close to the onset of the rainy season in this region and we can clearly see that hardly any net rainfall has occurred within the target area. In the surroundings however, rainfall amounts have been significant. The consequent storage

change may therefore have been picked up by GRACE and may have leaked into the solution for the upper Zambezi, both due to the limited spatial resolution of GRACE and the spatial filter applied. A longer filter correlation length results in further smoothing, which compromises the appearance of these details. Fig. 4.4 illustrates the concurrent 3-4
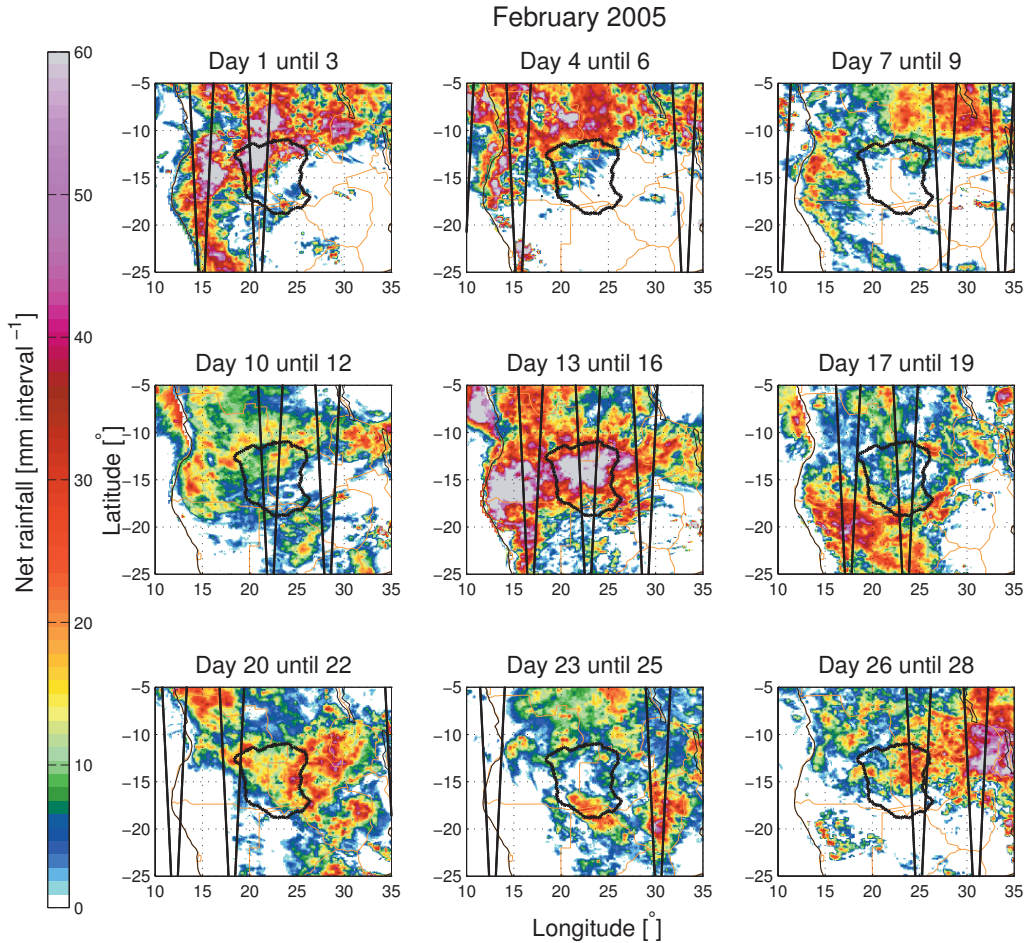
Figure 4.4: Ground track coverage during 9 intervals of 3-4 days in February 2005. The target area is delineated in black. Black dots indicate the ground tracks.

daily ground track patterns within February 2005, in which first of all the rainfall was irregularly distributed over the month and second, GRACE overpasses were not equally spaced in time. Most rainfall occurred within the first half of the month. There was quite a good GRACE sampling during these rainfall events. However, during the last half and especially the last week of February, GRACE did not cross the area in and around our

target area very often, implying that our target area has been poorly sampled within this period. The combination of more rainfall in the beginning of the month and poor coverage in the end of the month suggests that for this month the weight of both sampling and storage change was more towards the beginning of the month. Thus a time stamp shift towards the beginning of the month could be appropriate. Quantification of such a time stamp shift would require regional inversion of GRACE data, which has not been performed in this study.

These examples show that there is a trade-off in the chosen spatial filter. A short filter may reveal the effects of temporal inconsistencies due to the way in which GRACE samples a target area. A large filter may smooth such effects but this goes at the cost of perhaps hydrologically important information such as the amplitude of storage and interannual variability in storage patterns.

4.2 Discussion and proposal of framework

One of the targets of this thesis is to study the possible use of GRACE in hydrological modelling of river basins. The previous paragraphs illustrated the type of uncertainties and problems that we should expect when dealing with GRACE in hydrology. While the possible effect of aliasing may be emphasised by plotting the overpass tracks over the within-month net rainfall, it remains difficult to quantify, especially for hydrological scientists. However, the problem of leakage may be addressed by allowing hydrological knowledge to be iteratively integrated in the GRACE monthly solutions for a specific river basin. In this section, a framework is proposed, suggesting such an iterative modelling exercise, which may be followed to optimally profit from the information content of GRACE while taking care of the leakage errors.

From the previous sections, we may conclude that a framework has to deal with the following limitations:

- GRACE has a limited spatial resolution. Use of GRACE for calibration remains therefore difficult, because the spatial variability of hydrological processes and climatological inputs is too large to allow for a fully lumped hydrological model at the scale where GRACE observations become meaningful. Conversely, calibration of a hydrological model with spatially distributed parameters on a basin-integrated storage estimate would imminently lead to a great deal of equifinality.
- Leakage reduces amplitude and removes not only spatial, but also temporal (even inter-annual) variability. Reduction of information content due to leakage should be considered.

The first limitation leads to the fact that at the river basin scale, it is more straightforward to use GRACE as a means of validation of hydrological model behaviour, rather than calibration. On global scales, calibration efforts could be performed (see e.g. Werth *et al.*, 2009), although the user should bear in mind the risk of equifinality or overconditioning. The second condition leads to the conclusion that we need a prior estimate of the bias due to leakage, to make GRACE storage estimates more meaningful at the basin scale. Since leakage is strongly controlled by spatial and temporal variability of storage, we need prior information of this variability (i.e. mass variation inside and outside the basin of interest) to estimate leakage. In areas, isolated by ocean, this is straightforward, as the mass variability of the ocean has already been estimated and subtracted from the signal by means of an ocean background model. Velicogna and Wahr (2006) showed for instance that over Antarctica a constant bias correction gives a satisfactory result. However, when the target area is surrounded by land mass, more prior information on the water storage variability is needed. If this information is available, it may be filtered with the same spatial filter applied on the GRACE estimates to find a proxy of the errors caused by spatial filtering. Ironically, such prior knowledge requires a hydrological model. Since the aim of the thesis is to construct a hydrological model by using GRACE as complimentary data, this results in circular reasoning. In other words, GRACE could provide information to enhance hydrological models. To make GRACE estimates meaningful, we need to correct for leakage errors. For this, a hydrological model is required that we may only be able to validly construct if we can use GRACE to validate it.

The framework proposed here, in fact considers this circular reasoning as an opportunity. To construct a hydrological model, validated by GRACE, an iterative framework (displayed in Fig. 4.5) is in fact needed. The modelling part of this framework largely follows the suggestions made by Beven (2001), with the exception that validation and henceforth the decision for further iteration, may be carried out by employing GRACE, rather than field observations of discharge. It consists of the following iterative steps: (a) a perceptual and consequently conceptual model is required, based on prior beliefs of the dominant processes and how they are interlinked. This model should have a satisfactory spatial distribution and, other than with the traditional model calibration problem, should spatially be expanded beyond the region for which it is intended to be used (in this thesis, the upper Zambezi river basin) in order to be able to estimate the bias induced by leakage (note: for global or continental hydrological models, this is not an issue, as long as the modeller only considers validation of interior basins and not basins, close to the spatial boundary of the model). This conceptual model may be calibrated, using any data available, other than GRACE storage estimates; (b) The first approximation of the hydrological processes in and in the surroundings of the basin, i.e. the rainfall partition-

ing, stores, thresholds and dominant time scales and resulting fluxes, may henceforth be employed to make a spatially distributed and time-variable estimate of the storage in the river basin and its surroundings. This estimate may then be used to estimate the bias, caused by leakage, which henceforth leads to bias-corrected GRACE storage estimates; (c) The hydrological model outputs are consequently confronted with the corrected GRACE storage estimate and after statistical validation, visual inspection, or any other means of evaluation, the modeller may decide to improve the model structure. The quantification of the full uncertainties of GRACE data is not a straightforward task, because of the complexity of all the error sources (i.e. measurement noise, background models, aliasing and limited resolution). An important note here is that error assessments under the assumptions that geopotential variability, that cannot be explained by a constant annual cycle is due to errors (e.g. Wahr *et al*., 2006), is not valid since this jeopardises knowledge about hydrology, which confirms the existence of year-to-year variability in climate, and therefore in storage. It is therefore in this step up to the modeller to consider the softness of GRACE as a data source. Trivially, a wrong model structure and henceforth wrong modelled estimates of the storage variations will to a certain extent lead to wrongly predicted leakage induced bias. However, the bias estimate, relative to the storage, is a function of the spatial variability of storage, rather than of the storage itself, meaning that errors in the storage estimates will only be marginally transferred to the bias estimate, as long as the spatial variability of the storage is represented. Even when we consider that there are uncertainties in the bias estimate, the results of step (c) should lead to the conclusion that the model structure should be improved anyway. Consider for instance the extreme case, where a hydrological model would predict a completely constant storage in space and time. Bias estimates would naturally be zero and would not change GRACE estimates. Confrontation of the model with GRACE would immediately lead to rejection of the model. When the model does predict spatially and temporally variable storage, the error in bias estimates will only be a fraction of the error in the prediction of spatial variability by the model. For instance, let us assume that the real bias in a given month is 30% and the hydrological model with assumed correct rainfall input gives a 20% error in spatial variability (which is quite large). Then this would lead to a 6% persistence of this error in the final GRACE estimate. Hence, iteration should be performed until there is a coherence between the hydrological model and bias-corrected GRACE estimates. The assumptions in this approach are therefore more related to the quality of input and calibration data. These are assumed to be of reasonable quality at the temporal and spatial scale at which GRACE is used.

In the next chapter, the framework, presented in Fig. 4.5 will be employed in a case study on the upper Zambezi river basin and its surroundings.
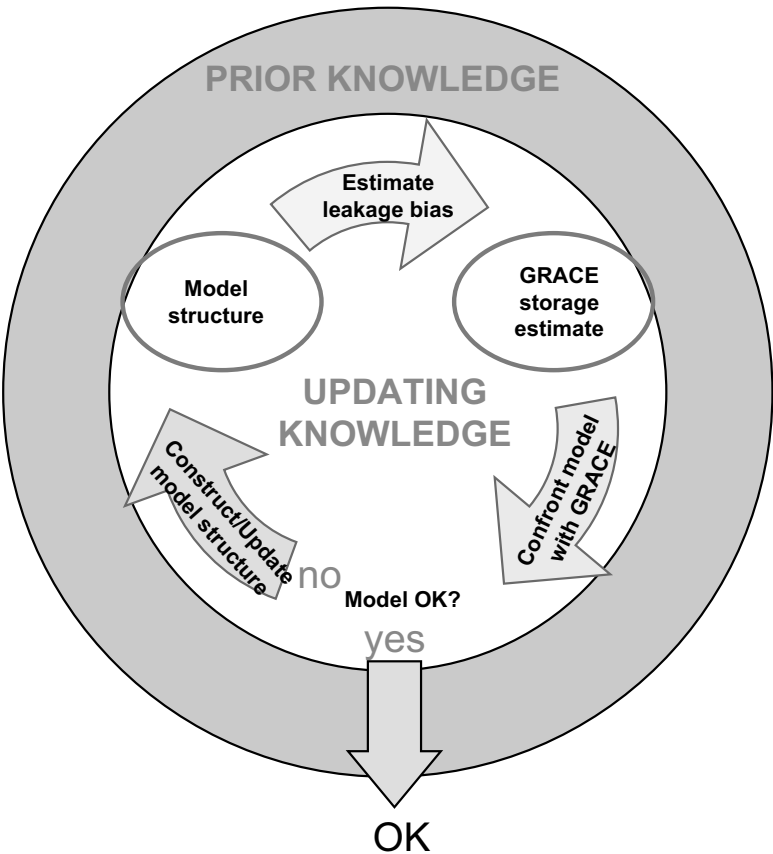
Figure 4.5: Framework to use GRACE as complementary information in the construction of hydrological models.

Chapter 5

Using GRACE in hydrological modelling of the upper Zambezi

This chapter deals with investigations on the use of GRACE data for the improvement and validation of hydrological models. The study follows the framework presented in Chapter 4 to construct a hydrological model of the upper Zambezi and its surroundings. First of all, an overview is given of how a short field survey has resulted in the formulation of the perceptual model of this area. Consequently, a conceptual hydrological model has been built based on this prior knowledge. Subsequently, the problem of GRACE bias due to leakage errors has been dealt with, by extending the hydrological model to the whole of Southern Africa. The month-to-month spatial variability of storage, estimated from this model, has been used to compute a time-variable bias estimate due to spatial filtering. Hence, GRACE could be corrected for this bias, after which it has been used for validation of the model structure and parameter values.

5.1 Introduction

Classically, a model architecture is estimated based on a perception of reality. Subsequently, parameters are calibrated based on discharge time series at the outlet of a catchment. Even in ungauged river basins, this strategy may be followed, when for instance an old discharge time series is available. The information content of such discharge time series is in many cases limited: the discharge may be available only at very large temporal (e.g. monthly) and spatial scales, which is often not appropriate for detailed modelling, or the

data may be of poor quality. Especially in distributed models, these problems generally result in "equifinality", where a large number of possible parameter sets performs equally well, but introduce high parameter uncertainty (Beven & Binley, 1992; Beven & Freer, 2001; Savenije, 2001). In the process of model parameter identification, it often appears that one poorly chosen parameter can easily "correct" for another poorly chosen parameter value, whereas physically these parameters are not correlated. Another problem with the calibration on discharge only, is that in many tropical regions, the discharge is often a relatively small flux compared to the rainfall. As a result, river discharge provides limited information on internal hydrological processes.

In order to prevent equifinality, a number of strategies may be followed: first, the modeller may limit the number of calibration parameters, which on the one hand may reduce equifinality, but on the other hand may jeopardise the physical representativeness of the model structure used; second, additional clearly identifiable information may be extracted from available discharge time series, other than residual time series (e.g. Winsemius *et al.*, under review, and Chapter 6); and third, the modeller may perform additional observations, that provide information on a different process in the water balance. Such additional observations, providing an independent perspective on a model's functioning, are here termed *orthogonal* observations. A frequently used term here is *multi-objective* optimisation. The author feels that this term should be reserved for calibration efforts, where a variety of criteria, derived from a single type of observation, is used. However, as soon as information from orthogonal data is introduced, a better reference would be *multi-observation optimisation*, because the hypothesis, encapsulated in the model is tested on "orthogonal" signals rather than discharge series alone. Evidently, such orthogonal observations may relate to changes in (certain compartments of) water storage and (certain sub-processes of) evaporation, given that the rainfall is known.

In this chapter, the focus is on the value of water storage observations in hydrological modelling. The aim is to use gravity observations from GRACE (see Chapters 3 and 4) as a means of hypothesis testing of a selected model structure for a hydrological model of the upper Zambezi river basin. GRACE provides in this sense "orthogonal" observations, meaning that it provides independent information on the dynamics of a certain variable within the model (in this case the storage), without being correlated to other sources of information, used for parameter or model structure estimation. These observations may help to gain insight into the validity of the selected model structure and parameter values, specifically those directly related to storage, meaning that it can be used to refine model structures and constrain parameters. A more plausible representation of storage within the model structure will lead to a more reliable simulation of the associated fluxes, which subsequently will result in more robust predictions, especially under conditions of future

change.

The goal of this chapter is reached by applying the framework, outlined in Section 4.2. A field survey of the upper Zambezi has been conducted to formulate a plausible perceptual model of the dominant processes. Henceforth, with the limited available ground data (described in Section 5.2), plausible model structures have been tested and rejected (Section 5.3). After this step, the model structure has been applied on the upper Zambezi and the surrounding river basins, which enables the use of this model as prior hydrological information for bias estimation, resulting from spatial smoothing of the GRACE estimates. Finally, the model structure and parameter values are evaluated with bias-corrected GRACE fields in Section 5.4.

5.2   Available data

At the time of study, data availability, both in terms of rainfall and discharge was limited. The most extensive dataset of monthly gauged rainfall was found in the database of the Global Historical Climate Network (GHCN) V. 2 (Vose *et al.*, 1992). A quite good coverage in the upper Zambezi was found in the period 1960–1972. The coverage deteriorated in the early seventies, especially over Angola, probably because of the struggle against Portuguese colonialism. The departure of the Portuguese and the start of the Angolan civil war put an end to most gauging activities, which means that a large part of the Zambezi upstream of Lukulu remained ungauged during this period. Spatial estimates of monthly rainfall were determined from GHCN stations for the period 1960–1972 by means of the weighted inverse distance method.

Unfortunately, some of the available discharge time series did not cover the same period as the rainfall data. Thus more recent rainfall records were needed. It was chosen to use the Climate Research Unit's (CRU) V. 2.0 monthly grids (New *et al.*, 2002). The grids $(0.5 \times 0.5 \ (\text{km})^2)$ are based on as many rainfall stations as could be found. It is to be expected that rain gauge coverage was in any way poor in the period where CRU fields were used for calibration. Therefore, this dataset was used to model the general behaviour of a watershed but only marginally to mimic details in the hydrograph. For running the models in recent periods, satellite rainfall from the Famine Early Warning System (see Section 3.2) has been used.

To obtain potential evaporation estimates, Penman-Monteith was applied on reanalysis data of NCEP/NCAR. It is likely that this approach only provides coarse estimates, but this is expected to be appropriate because soil moisture is in most places the limiting factor for evaporation, except during the wet season in flooded wetlands.

Monthly discharge records were available for a small number of gauging stations within the upper Zambezi. The most important stations were located in Kabompo, Lukulu and Victoria Falls (see Fig. 5.1). Furthermore, for the calibration of models, surrounding the upper Zambezi, all available discharge data for this region were taken from the database of the Global Runoff Data Centre (GRDC, see `http://grdc.bafg.de/`). The spatial distribution of discharge stations is presented in Fig. 5.10.

5.3   Development of a monthly scale hydrological model

5.3.1 Field survey of the upper Zambezi: from the source to the Victoria Falls

The Zambezi river springs from the northern areas of Zambia in the Kalene Hills, close to the border with Congo. From there, it flows to the west into Angola and shortly afterwards bends southwards entering the Western Province in Zambia. The surroundings of many of the Northern tributaries are confined by seasonal floodplains and are governed by mild slopes, consisting of highly permeable sandy soils, deposited in the tertiary and quarternary age as an outcrop of the Kalahari desert. As a result, infiltration capacity is high. In combination with relatively high annual rainfall rates in the North, this results in a high availability of soil moisture in the wet season, which facilitates the existence of lush (partly ever)green vegetation (also shown in the photo on Fig. 5.2).

In October 2005, a field survey of the upper Zambezi was conducted, which is here defined as the area upstream of Victoria Falls (ca. 500 000 (km)^2). An overview of the surveyed route is given in Fig. 5.1. This field survey was useful to collect 'prior knowledge', indicated in Fig. 4.5, and in this case, this knowledge has been used to construct a perception of the catchment architecture, which was subsequently used to set up and code a hydrological model.

The mild slopes in the terrain have in many places led to the evolution of small groundwater dominated wetlands within the river system, locally known as 'dambos'. Most dambos in fact seem to connect river stretches of lower order streams and act as a fill-and-spill system, meaning that they cause a delay if a flood wave occurs. All dambos together are able to store large amounts of streamflow over longer periods, which eventually evaporate (see also Bastiaansen, 1995). An impression of these wetlands is given in Fig. 5.3, which is a photograph of a typical dambo, very nearby the Zambezi river. In many dambos, scoop holes were found, giving the opportunity to measure the local groundwater level. It turned out that in the end of the dry season, groundwater levels in the dambos were
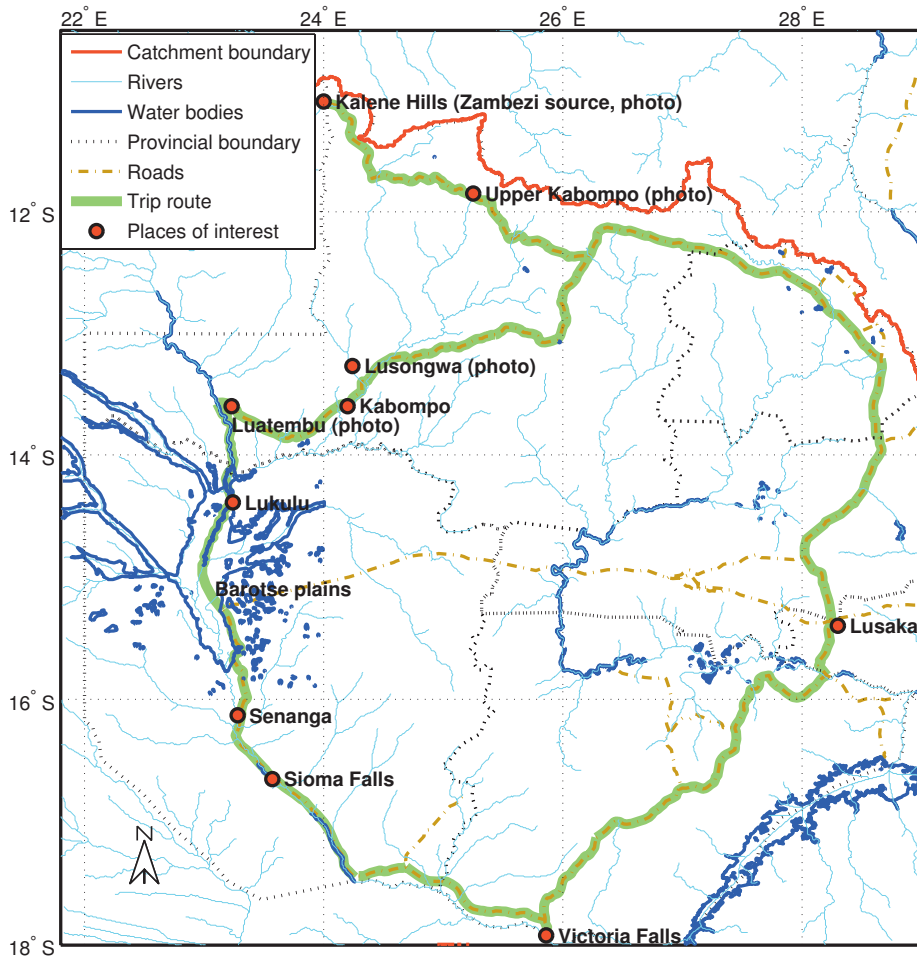
Figure 5.1: Map of field work route, taken during the field trip for this thesis in October 2005. Places of interest from which a picture has been included have been additionally marked with 'photo'.

mostly less than 1 meter below surface level (for instance in the Lusongwa plain, Fig. 5.4.

Due to the deep layers of deposited permeable sand, it is believed that most of the discharge in these rivers is generated from groundwater. This hypothesis is further supported by the relatively high discharges in the more upstream low order rivers in the end of the dry season, where the influence of the many fill-and-spill dambos is still small, and the remarkable low turbidity of the water. During our field survey, a photograph of one of the low order streams was taken, close to the border of the Kabompo catchment (Fig. 5.5).

Figure 5.2: Kalene Hills, the source of the Zambezi. Here, the mighty Zambezi starts its long journey from the North of Zambia, eventually entering the Indian ocean in Mozambique. Here, the equatorial vegetation is able to sustain leaves throughout the whole year. *left*: Michael Mutale, field expert, *right*: Hessel Winsemius, author of this thesis.

Upstream of the wide Barotse floodplains (in Fig. 5.6 bordered by Lukulu in the North, Mongu in the East and Senanga in the South), a number of important tributaries such as the Kabompo, Luena and Luanginga join the Zambezi. The river converges at Lukulu before spilling into the Barotse plains: a shallow wide floodplain area consisting of several tens of meters deep Kalahari sands, an enormous phreatic groundwater reservoir. This area is completely flooded during the peak of the wet season, but remains completely dry in the remaining months (see Fig. 5.6, where this is envisaged with satellite imagery from the end and the beginning of the dry season). More downstream, the river bends to the west where it drops down the Victoria Falls. Elevation maps suggest that the Okavango and Cuando rivers enter the Zambezi east of Victoria Falls through the Caprivi Strip. However, flow through the Caprivi Strip has only been known to occur during extremely wet years.

Figure 5.3: Dambo, crossed by the Luatembu river in Northern Province, Zambia.

5.3.2 Model perception

The field survey resulted in a number of conclusions about the dominant hydrological processes, which are summarised below:

- The upper Zambezi catchment is an outcrop of the Kalahari desert, and therefore consists of deep layers of permeable sand, deposited by wind. While rainfall amounts in the Kalahari, South of the Zambezi river, are too small to generate considerable recharge (De Vries *et al.*, 2000), in the upper Zambezi, a considerable amount of baseflow is sustained all through the year. The permeability of the soil and the mild slopes do not allow much overland flow, which means that in fact almost all discharge is generated from groundwater (also proven by the low turbidity of the water).

- In the upper reaches, the discharge regime is governed by a system of wetlands (dambos). They are able to store a considerable amount of the discharge, generated from the first rains, and act as a fill-and-spill system (equivalent to a Nash cascade system, with an upper threshold and probably very little release through

Figure 5.4: A scoop hole in the Lusongwa plains, a dambo near the town Zambezi, in the Kabompo catchment.

the bottom). This system results in the occurrence of a regime switch when the drainage through interconnection of dambos is activated. A lot of the discharge spills into the dambos and is eventually lost in evaporation (either from open water or from the neighbouring woodlands), which results in attenuation of the downstream flow.

- The more downstream reaches are bordered by large seasonal floodplains (the largest being the Barotse floodplains) which are again able to store and evaporate a considerable amount of discharge in the permeable and very deep sand bed. The result is that there is a considerable time lag of up to three months between the peak of the hydrograph at Victoria Falls and the peak of the rainy season of the upstream area.

- Vegetation in the Northern parts consists of dense, partly evergreen woodlands, that can be sustained because of the high water retaining capacity of the area, combined with high rainfall and infiltration rates.

Figure 5.5: The upper Kabompo river, one of the major tributaries of the Zambezi, close to its origin (see Fig. 5.1). These tributaries sustain a great deal of base flow throughout the dry season and have remarkably low turbidity levels.

A schematic overview of these hydrological processes (i.e. the perceptual model) is shown in Fig. 5.7.

In the uplands, rainfall is partitioned into direct evaporation from the surface (interception) and infiltration into the unsaturated zone, where it partly transpires and partly percolates. It is expected that surface runoff only occurs due to saturation of the soil in the lower lying areas and wetlands (dambos). In the wet season, the groundwater levels
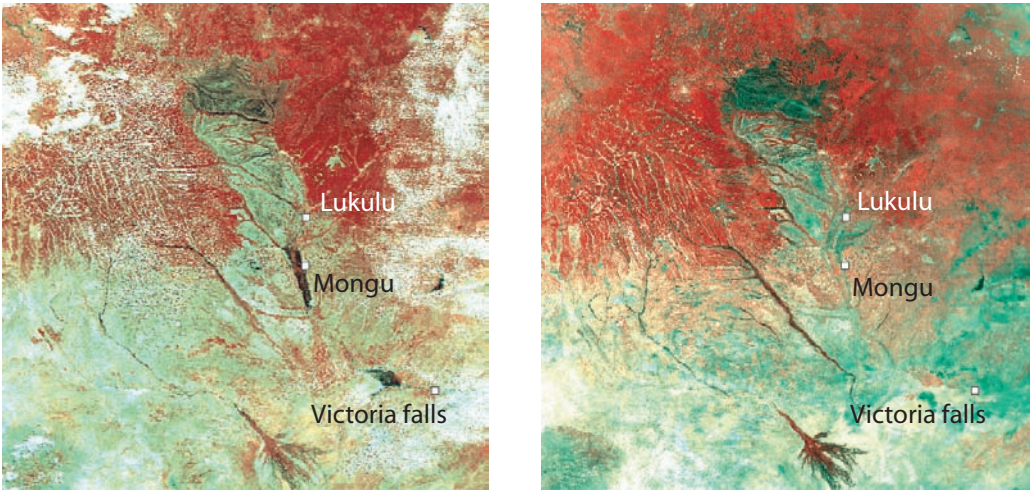
Figure 5.6: MODIS 250m channel 2, false colour images of the study area. Left: the beginning of the dry season. Right: the end of the dry season. Dark pixels indicate flooded areas and are concentrated around major tributaries in the North. Red pixels are forested by equatorial species.
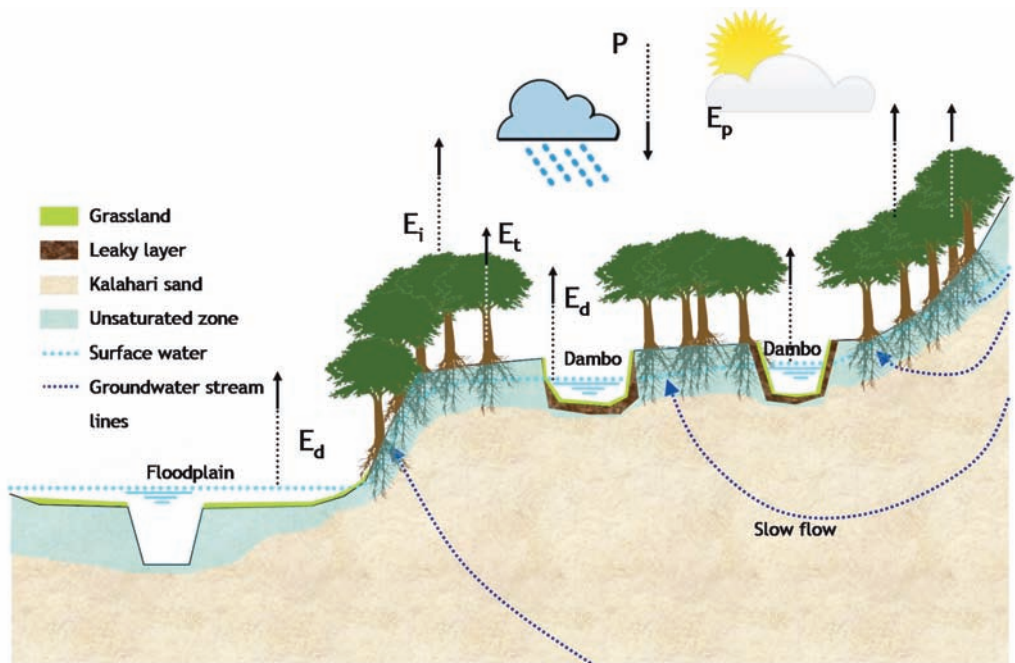
Figure 5.7: Model perception of the upper Zambezi.

rise and the dambos become flooded. Quick response flows overland are generated by threshold behaviour within the channel-dambo system: when the dambo levels exceed a threshold level, the drainage network becomes interconnected. Slow base flow is generated from the deep groundwater and is therefore sustained all through the year. In the flood-plains, the low slopes have resulted in a low channel capacity. In the flooding season, this results in overland flow which means that flood peaks are further attenuated by spillage into floodplains (the largest being the ∼ 20 km wide Barotse plains, also indicated in Fig. 5.1, consisting of deep Kalahari sand deposits). No major changes in the hydrology or land cover of the upper Zambezi have taken place over the past century. Therefore, it is assumed that the river's hydrological behaviour may be considered to be stationary.

5.3.3 Model setup and calibration

The model has to meet a number of requirements:

- The number of parameters should be small, preferably limited to the ones that may be estimated using available data.
- The conceptual model structure should clearly represent the natural behaviour of the river basin and,
- the spatial variability of dominant hydrological processes should be encapsulated in the model structure, which may require more parameters than parsimony would allow and thus may conflict with the first criterion.
- The time step of the model should be equal to or smaller than one month, first of all since dominant processes are expected to have a time scale of this order of magnitude (except for evaporation from interception) and second because GRACE typically is processed on monthly time scale.

It is clear that to some degree, there will be a trade-off between criteria. It was decided to construct a monthly time step semi-distributed conceptual model, consisting of 'Lumped Elementary Watersheds' (LEW). LEWs are spatial model units that are treated lumped, but may possibly be modelled by means of different conceptual model structures, suitable for the specific dominant hydrological processes (also called hydrological response units (HRU) in other studies). The outlets of individual model units may be located at gauging points, confluence points with a higher order stream or at any point where a transition between different dominant hydrological processes is observed.

On the basis of the perceptual model, several model structures have been set up and evaluated. To this end, the CRU rainfall database has served as input while old river discharge data (1992–2001) from the Kabompo river, measured near Kabompo (see Fig. 5.1) have been used for evaluation purposes. The hydrological response of the Kabompo

is assumed to be representative for the response of all upper reaches. Several model structures, all encapsulating the prior knowledge collected in the field, were iteratively tested. This was done by estimating some parameters a priori, some by visual inspection of the hydrograph and some (i.e. the ones not directly identifiable from signatures in the data) by performing calibration by Monte-Carlo sampling. To assess model performance, the coefficient of Nash and Sutcliffe (1970) has been used on the residual discharge series of the Kabompo, but also on the residual of the log of discharge time series over the available period 1993–2001. Eventually, a model structure has been chosen that yielded acceptable identifiability of parameters, subject to calibration, while maintaining parsimony and representativeness of the hydrological processes. The final result is schematically displayed in Fig. 5.8 and described below.

5.3.4 Conceptual model

Net precipitation $P_n(t)$ is calculated by subtracting a fast evaporation threshold, D [L T^{-1}] as suggested by Savenije (1997, 2004). The fast evaporation consists of interception (evaporation of rainfall within a day after the event took place) and evaporation from the top soil layer, including transpiration from shallow-routing vegetation that transpires within the monthly time step. For convenience this fast evaporation is called interception $E_i(t)$. E_i and P_n are computed by

$$E_i(t) = \min\left[P(t), D, E_p(t)\right] \tag{5.1}$$

and

$$P_n(t) = P(t) - E_i(t). \tag{5.2}$$

The remainder of the model structure consists of two spatially defined zones: the first is the highland zone, a combination of two reservoirs representing the unsaturated ($S_u(t)$ [L]) and saturated zone ($S_s(t)$ [L]), while the second consists of one reservoir representing the wetland zone (S_d [L], where the subscript d stands for 'dambo'). This zone accounts for the behaviour of dambos and floodplains. Both zones account for a part of the total watershed drainage area, determined by a surface fraction A_{wet} [-], which may be estimated roughly using satellite imagery. The zones are interacting by an overtop of the saturated zone, which occurs when S_s exceeds $S_{s,th}$ [L]. The unsaturated zone is modelled similar to the unsaturated zone in the HYMOD model (Boyle, 2000; Vrugt et $al.$, 2003): in this structure, it is assumed that the spatial variability of soil moisture capacity S_{max} [L] can be described by a power function representing the spatial variability of what is, in fact a threshold process,

$$F\left[S_u(t)\right] = 1 - \left[1 - \frac{S_u(t)}{S_{max}}\right]^B, \tag{5.3}$$

Figure 5.8: LEW model structure of the upstream catchments of the upper Zambezi.

where $F(t)$ [-] represents the saturated fraction of the modelled drainage area and B [-] determines the spatial variability of the soil moisture capacity. $P_{e1}(t)$ and $P_{e2}(t)$ [L T^{-1}] are excess rainfall components that are partitioned according to the moisture state of the catchment. When a fraction of the basin $F(t)$ is saturated, P_{e2} will only be generated over this fractional area. P_{e1} is the excess rainfall, occurring when $F = 1$. Transpiration from S_u is described according to the following relationship:

$$E_t(t) = \min\left(\frac{S_u(t)}{l_p S_{max}}, 1\right) E_{pt}(t),$$ (5.4)

where $E_t(t)$ [L T^{-1}] is the actual transpiration, l_p [-] is the fraction of available soil moisture where vegetation becomes moisture stressed, here fixed on 0.5 according to Rijtema and Aboukhaled (1975). S_{max} [L] is the maximum soil water holding capacity in the root zone and $E_{pt}(t)$ [L T^{-1}] is the potential transpiration (i.e. the energy available for transpiration,

$E_p(t) - E_i(t)$). Capillary rise ($C(t)$ [L T^{-1}]) is considered dependent on the amount of available soil moisture according to the relationship

$$C(t) = \min \left[C_p \frac{S_{max} - S_u(t)}{S_{max}}, S_s(t) \right], \tag{5.5}$$

where C_p [L T^{-1}] is the maximum capillary rise. Flooding of the dambos occurs when overtop of the saturated zone occurs. Excess water spills into the wetland zone where it is subject to open water evaporation and transpiration by shallow rooted vegetation $E_d(t)$ [L T^{-1}] (see Fig. 5.3) and two runoff mechanisms for surface and groundwater drainage (K_s and K_q [T]), behaving like linear reservoirs. K_q represents the drainage network between the wetlands. This network only generates flow when a threshold $S_{d,th}$ [L] is exceeded. K_s is not dependent on the river network and should therefore be equal in the highland zone and in the wetland zone. River routing is simulated by using lag times, estimated from available downstream discharge records at Lukulu and Victoria Falls.

A number of parameters can be estimated directly from the recession periods in the hydrographs. Visualisation of the natural logarithm of discharge provides a good indication of K_s, K_q and $S_{s,th}$. The first two parameters express themselves through the slopes in the recession curve, the latter is found at the inflection point of two subsequent slopes. Some visual calibration was necessary to correct for second order effects like capillary rise and evaporation. It was therefore decided to fix D and C_p for most target areas. A_{wet} has been estimated from the MODIS satellite images, shown before, S_{max}, B and $S_{d,th}$ could not be visually identified and have therefore been included in the Monte-Carlo simulations. Fig. 5.9 shows a scatter plot with the three selected parameters displayed on the x-axis and the Nash-Sutcliffe coefficient for low and high flows displayed on the y-axis. The posterior parameter structure demonstrates identifiability. Particularly B and $S_{d,th}$, responsible for the areal distribution of soil depth and the dambo interconnectivity threshold, appear to be significantly identifiable, the latter predominantly in the low-flow criterion. It is interesting to note that the likely value of $S_{d,th}$ of 35 mm could be equivalent to the groundwater level under surface level, observed in the field. Given a storativity of about 10% in the clay deposits in the dambos, a threshold of 35 mm before flooding occurs would mean a depth of the groundwater of $35/0.1 = 350$ mm which is in the order of magnitude of the observation in the field (groundwater in dambos less than 1 meter below the surface). It is likely that near the stream, crossing the dambo, this threshold is smaller than further away from the stream. Finally, the sensitivity of S_{max} towards discharge reduces considerably with larger values. Therefore it was decided to fix this parameter at a value, slightly higher than the annual rainfall. It is expected that storage behaviour will also be insensitive to any higher selected value for S_{max}. The model structure for individual model units does not yet take care of the possibility of overbank
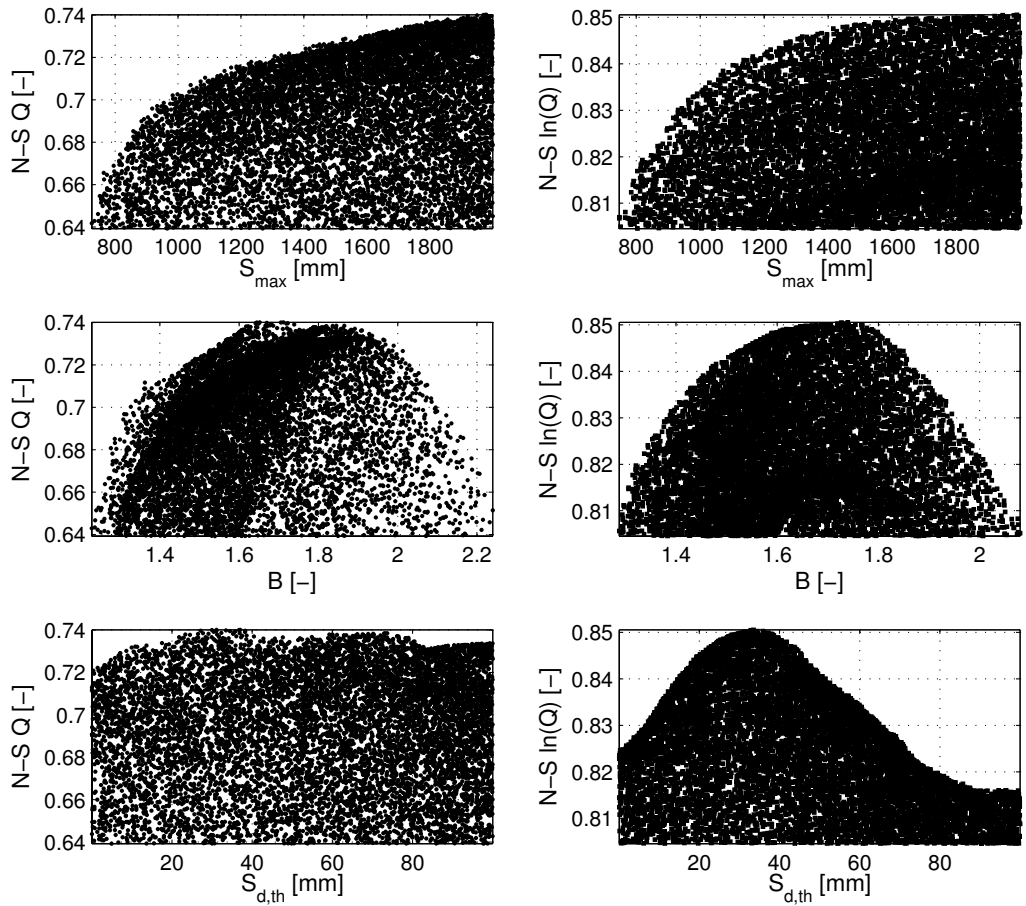
Figure 5.9: Scatter plot of Monte-Carlo results.

discharge. Therefore, a threshold Q_{max} [L T^{-3}] has been introduced, which separates river flow from overbank flow. Since Q_{max} occurs downstream of the model structure presented in Fig. 5.8, it is not part of the calibration procedure.

Most of the parameter values, derived for the Kabompo river have been adopted as well for the other tributaries, upstream of Lukulu, where discharge data was not available. This was done, because only at Lukulu, a second gauging station is present. Q_{max} has been fixed per watershed. It determines to a large extent the peak behaviour of the hydrograph in the Zambezi. Peaks are seriously attenuated when low values for Q_{max} are chosen and therefore the degree of attenuation between subsequent stream flow time series has been used to tune this parameter. The Barotse floodplain bordering the Zambezi has been considered independently from the upstream catchments. The parameter A_{wet} is larger for

the more downstream model units, since these areas contain a much larger surface, taken by dambos and floodplains. Downstream of Lukulu, the same approach has been followed: most parameters were set at equal values for all watersheds except for Q_{max} and A_{wet}, because Victoria Falls was the only remaining discharge station, that could be used for calibration. Fixing parameters over large areas prevents over-parametrisation. This means that most of the areal distribution of stores and fluxes is caused by the distribution of the rainfall input. Table 5.1 shows the parameter set obtained.

Table 5.1: LEW parameters. In ungauged model units, the values were adopted from gauged neighbouring catchments as much as possible. (e): estimated by expert knowledge, (v): estimated from visual inspection of discharge records, (c): calibrated by Monte Carlo sampling.

Sub-catchment	D (e)	S_{max}(c)	B (c)	C_p (e)	$S_{s,th}$ (v)
	[mm month^{-1}]	[mm]	[-]	[mm month^{-1}]	[mm]
Highlands Lukulu	50	1600	1.65	2	50
Wetlands upstream Lukulu	50	1600	1.65	2	120
Luanginga	70	1000	1.8	2	100
Upstream of Victoria Falls	70	1000	1.8	2	100
Barotse plains and surr.	70	1000	1.8	2	100
	K_s (v)	$S_{d,th}$ (c)	K_q (v)	A_{wet} (e)	
	[month]	[mm]	[month]	[-]	
Highlands Lukulu	8	35	4	0.1	
Wetlands upstream Lukulu	12	35	8	0.3	
Luanginga	3	35	3.5	0.1	
Upstream of Victoria Falls	3	35	1.2	0.1	
Barotse plains and surroundings	3	35	1.2	0.7	

5.4 Evaluation of the LEW model with GRACE gravity fields

5.4.1 How can prior hydrological knowledge serve as bias estimate for GRACE?

The previous modelling attempt is purely based on expert knowledge and some ground data and it only covers the upper Zambezi. In order to retrieve meaningful signatures from GRACE (i.e. the right annual amplitude, the right inter-annual variability) a bias correction needs to be applied. In Chapter 4, it is hypothesised that bias estimates may be derived from prior knowledge of the spatio-temporal storage variability. Formulated, the temporally variable basin-averaged storage, equivalent to what we expect GRACE to

measure may be computed from the hydrological model by

$$S_A(t) = \int\limits_{A_\xi} \left[S(A,t) - \overline{S}(A) \right] \mathrm{d}A, \tag{5.6}$$

where S_A is the basin-averaged storage of target area A_ξ, $S(A,t)$ and $\overline{S}(A)$ are the spatio-temporal variable storage and the spatially variable long-term mean storage, both taken from the LEW hydrological model and A_ξ represents the spatial extent of the target area. An effective bias estimate (i.e. the combined effect of signal, leaking into and out of the area of interest) can then straightforwardly be computed by

$$\epsilon_A = S_A(t) - \int\limits_{A_\xi} \left[S(A,t) - \overline{S}(A) \right] * f(A)\,\mathrm{d}A, \tag{5.7}$$

where $f(A)$ is the applied spatial filter (see also Eq. 4.1). Eq. 5.7 can only be applied when the spatial extent of the hydrological model exceeds the effective area of the spatial filter, applied to the target area under consideration.

5.4.2 Extention of LEW to Southern African river basins

In order to fulfil this condition, a considerable spatial expansion of the LEW model is required. For this study, many of the surrounding basins have been considered, in particular the Shebelle, the upper Nile, Congo, Zambezi, Okavango, Limpopo, Orange and many minor coastal basins. For these basins and in most model units, the same model structure as presented in Section 5.3.3 has been applied, even though the correctness of the model structure of the upper Zambezi is not guaranteed for all the surrounding basins. Fig. 5.10 shows the river basins, considered in the model and the individual model units. This target area is far greater than needed for validation of the upper Zambezi, but in this study, it was decided to try to evaluate GRACE over a multitude of target areas in order to make the conclusions more robust and to investigate the relation between size of the area and the information content of GRACE as well as the contribution of the bias. A potential contribution to the bias of water storage variations outside this area has been neglected. This is justified because (i) GFZ has already reduced the contribution of oceans by means of an ocean model prior to the estimation of monthly GRACE models; (ii) the continental areas are generally outside the significant support of the filter function for the chosen target areas.

Within the major basins, many model units have been delineated. Major lakes and wetlands such as Lake Kariba, Lake Cahora-Bassa and Lake Malawi, have been delineated as
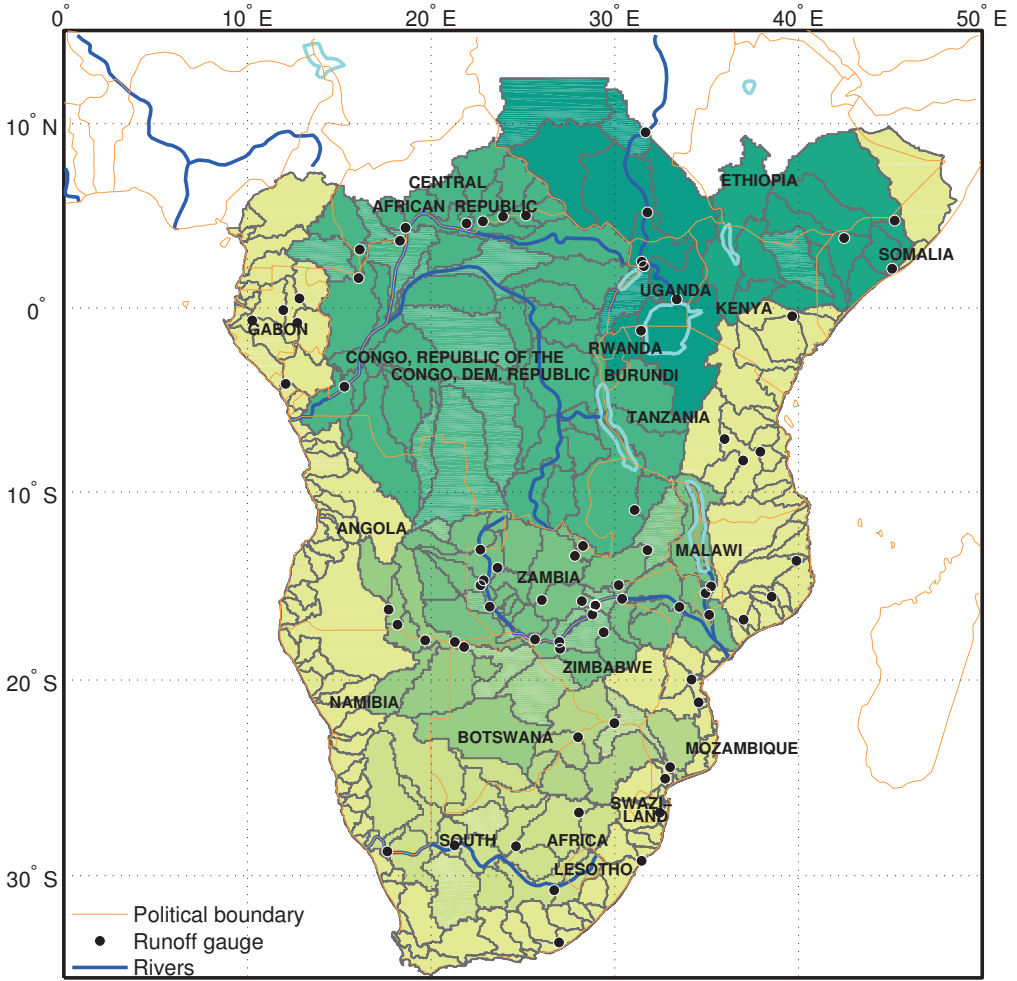
Figure 5.10: Geographical extent of the LEW model for Southern Africa. The grey lines represent the delineations of model units. Large rivers and lakes are indicated in respectively dark and light blue. The major river basins selected for this model are shown in a colour gradient from dark-green to yellow. The black dots indicate locations of runoff gauges, from which monthly stream flow records are available and have been used for calibration.

separate model units and their storage variability has been modelled with simple water balance equations. The water balance of a lake has been modelled as

$$\frac{\mathrm{d}S}{\mathrm{d}t} = Q_{up}(t) + [P(t) - E_p(t)] A_l - Q_{down}(t), \tag{5.8}$$

where A_l [L^2] is the total lake surface (assumed to be constant), Q_{up} is the inflow into the lake from upstream and Q_{down} [L T^{-3}] is the outflow from the lake. For natural lakes, the outflow has been estimated as

$$Q_{down}(t) = a \max(h(t), 0)^b, \qquad (5.9)$$

where a [L^{3-b} T^{-1}] is a calibration rating constant and b [-], a rating constant, assumed to be equal to 1.5. h [L] is the water level in the lake with respect to a reference level, being the no-flow threshold of the downstream river. For man-made lakes, storage and releases have been modelled by means of simple operating rules, where outflow is equal to a draft flow, except under extreme conditions such as floods or droughts. The water balance relations were roughly calibrated on sparsely available water level information from altimetry missions, such as TOPEX/POSEIDON and Jason-1, distributed through the ESA 'River and Lake' project (`http://tethys.eaprs.cse.dmu.ac.uk/`). The flow behaviour in the upper Nile is particularly dependent on the water level in lake Victoria. Therefore, a separate LEW has been defined for lake Victoria, modelled as described above. In the Okavango river basin, the Okavango delta has been delineated manually to take into account the surface runoff redistribution which is actually taking place in this delta. The Okavango's neighbouring interior basins have been modelled as a purely vertically flowing system, meaning that there is no lateral exchange of water between these basins and no significant runoff takes place. The same holds for the two most eastern model units of the Shebelle. These basins drain on the salt lake Turkana (North-Kenya Rift Valley), from which no runoff flows to neighbouring catchments.

5.4.3 Regional water management and wadis

A complication of large-scale water balance modelling is the regional lateral interactions that take place between river and surrounding areas, either because of human interference or topography, geology and climatology. Large dams and lakes may be relatively simple to include in the LEW modelling approach, but information is needed about their operation and general behaviour (e.g. surface–volume relations and operating rules) to adequately model their water balance. For the application that is presented in this chapter, a rough estimate of their water balance should be adequate. In addition, many irrigation schemes are present in Southern Africa. Some of the largest irrigation developments are taking place in the nearby Limpopo river basin (South of the Zambezi). But also in the Orange basin, the lower Zambezi and the Incomati, some irrigation is taking place. Also wadis are present in the modelled area, e.g. in the Shebelle river basin, where runoff is generated in the Ethiopian highlands, which ends up in the downstream desert area. The redistribution of surface runoff over either irrigated areas or wadis has been included by spilling a certain

amount of runoff in model units that are suspected to be interspersed by irrigated areas, wadis or wetlands. Per time step, this runoff is simply added to the total rainfall in the downstream neighbouring model unit. For other applications where these interactions are of greater importance, more details about these interactions must be included in the model.

5.4.4 Outputs concomitant with GRACE time series

LEW water storage estimates, concomitant with GRACE time series, have been generated using rainfall estimates from the Famine Early Warning System (FEWS, see Section 3.2). These estimates have been lumped over the model units to provide a time series from January 2001 until June 2006. The first two years of simulation have been used as spin-up time to stabilise the state variables of the LEW model structures, which leaves the GRACE period January 2003 until June 2006 for evaluation.

5.5   Results and discussion of effectiveness

The magnitude of the extended LEW model, allows the analysis of target areas of different sizes, in the range of 10^5–10^7 (km)^2. Therefore, four target areas were defined (Fig. 5.11): 1. upper Zambezi (UZ), the focus of this study, with an area of 4.7×10^5 (km)^2; 2. Zambezi (Z), 1.3×10^6 (km)^2; 3. upper Zambezi + Okavango (UZO), 1.2×10^6 (km)^2; 4. Zambezi + Congo (ZC), 5.2×10^6 (km)^2. Gaussian spatial filters with correlation lengths of 600 km, 800 km, and 1000 km have been used and the bias has been estimated for each filter.

5.5.1 Bias, estimated from LEW model

Monthly bias estimates have been computed from the output of LEW using Eq. 5.7. The computations have been done for each target area and choice of Gaussian filter correlation length. Figure 5.12 shows as an example the bias time series for the four target areas using a 1000 km Gaussian filter. A number of observations can already be made:

- The bias-to-signal-ratio is significant. A comparison of Fig. 5.12 with Fig. 5.13 reveals that the annual amplitude of the bias may even exceed the amplitude of the water storage variations (i.e. the signal). This underlines the need to correct GRACE estimates of monthly mean water storage variations for the bias introduced by spatial smoothing, prior to its use for hydrological model evaluation. As expected, the bias strongly depends on the correlation length of the filter.
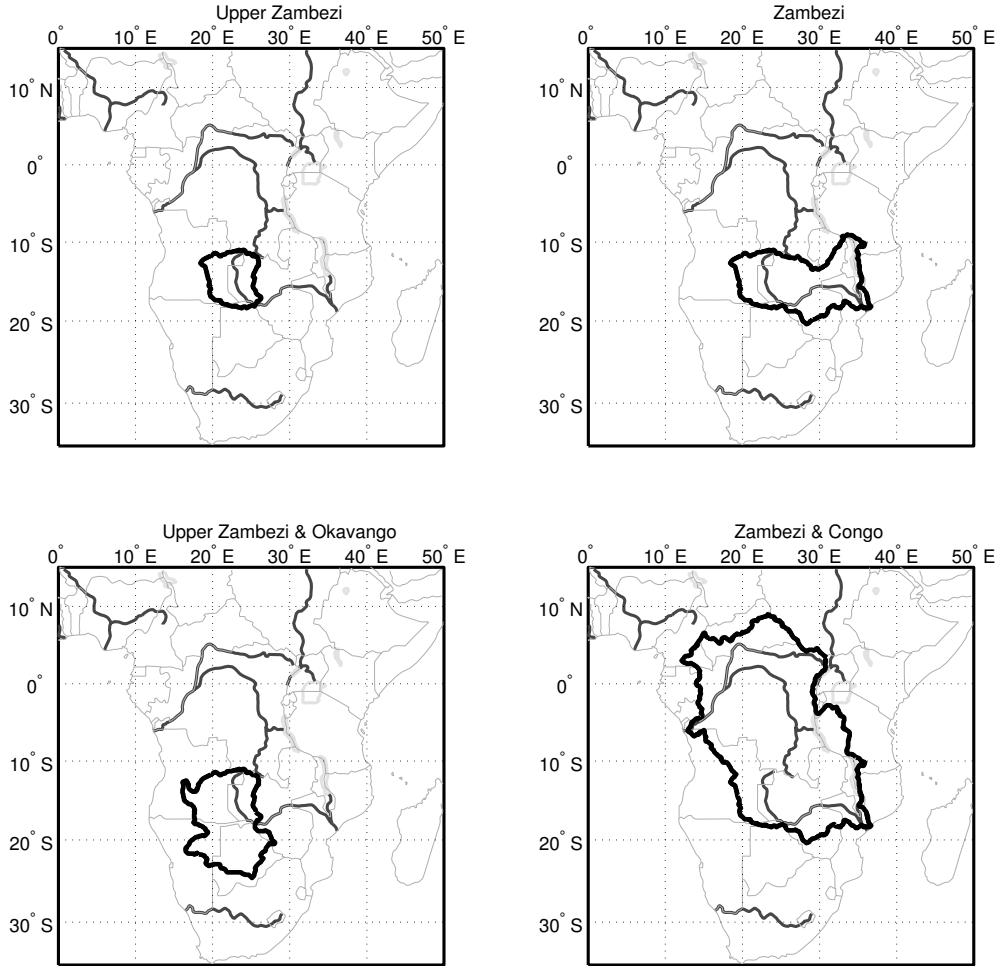
Figure 5.11: The four target areas, used in this study.

- There is a strong periodicity, and, more importantly, there is a great deal of inter-annual variability in the bias signal. This implies that a spatial filter may in fact remove a great deal of information from GRACE data, encompassed in its interannual details. The reason for this is related to the typical monsoon patterns, which are the cause of large spatial gradients in rainfall, and consequently, storage. The interannual variability also underlines that in some years, this spatial gradient may in fact be smaller than in other years (i.e. the rainfall is more evenly distributed in space than in other years).

- The bias and the bias-to-signal ratio depend on the size of the target area: the
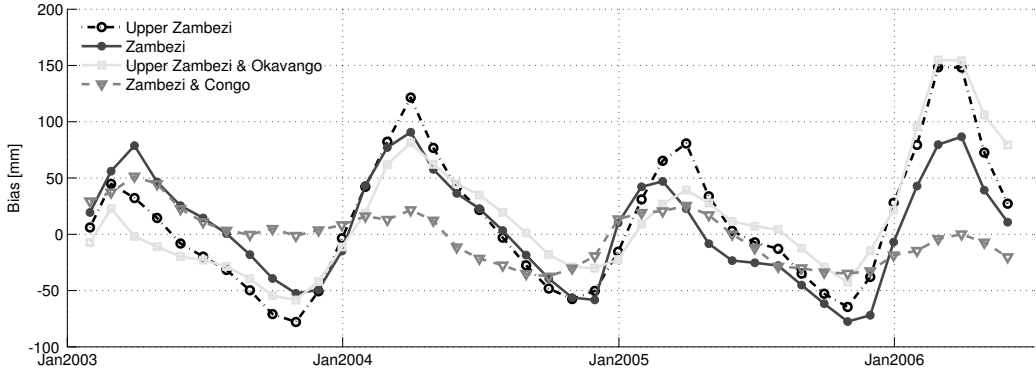
Figure 5.12: Time variable bias for the target areas using a gaussian smoothing function with correlation length of 1000 km.

smaller the target area, the larger the bias and the bias-to-signal ratio. For instance, for the smallest target area, the upper Zambezi (4.7×10^5 (km)^2), the amplitude of the annual bias is 47 % of the annual water storage variation when 1000 km Gaussian smoothing is applied. For the largest area, the Zambezi + Congo, (area 5.2×10^6 (km)^2), the annual bias reduces to about 31% of the annual water storage variation. This is first of all due to the effect of the spatial filter applied. In relative sense, its effect becomes smaller with increasing area. Secondly, monsoonal differences are smaller over the larger target areas. The Congo has in fact a bi-modal rainy season. This not only reduces the amplitude of storage, but also the spatial variability in storage, thus reducing the bias.

The estimated bias has been used to correct GRACE monthly mean water storage variations. Here, 31 monthly GRACE gravity field models between January 2003 and March 2006 were used from GFZ's release 03.[*] For more general information about GRACE, the reader is referred to Section 3.1. From the time series of monthly gravity field models, monthly water storage variations have been computed following the procedure of Swenson and Wahr (2002). These estimates have been corrected for the bias. For this purpose, the bias estimates, computed from the LEW model output, have been spline interpolated to the time stamps of the monthly GRACE models.

Fig. 5.13 shows the time series of monthly mean water storage variations. It is clear that the bias correction has a substantial impact, particularly in the smallest target area, and that the correspondence between bias-corrected GRACE and LEW is quite good,

---

[*]. Note that since the publication of this work in Klees *et al.* (2007), many work has been done on improvement of gravity field retrieval and new filters have been developed, which increase the signal-to-noise ratio. This is further discussed in Section 5.6.3.
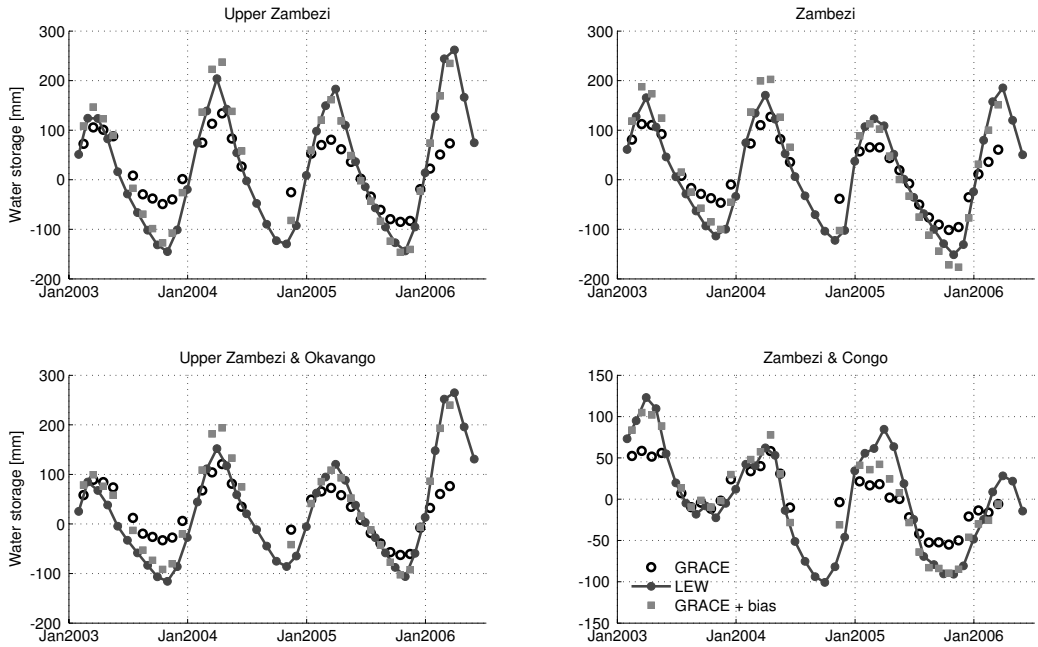
Figure 5.13: Time variable storage for different target areas with gaussian filter, correlation length 1000 km: The black circles represent GRACE without bias correction (**GRACE**), the blue lines with dots represent simulated storage by the LEW model (**LEW**) and the red dots represent GRACE with bias correction (**GRACE + bias**). Note the different scale of the y-axis in the results for the Zambezi and Congo target area.

especially in the upper Zambezi, the focus of this study. In fact, all target areas show a consistent storage pattern with the same order of magnitude in terms of annual amplitude. A remarkable result is that the interannual variability of the storage in all catchments is consistent. High peaks are reproduced in both data sets while the storage in the end of the dry season is also consistently reproduced. Especially in the Zambezi + Congo target, the end of the dry season reveals a trend which is shown in both GRACE and the hydrological model. Such trends are significant hydrological information, which was also shown by (Chambers, 2009), who used prior knowledge of the natural variability of land water storage by a hydrological model, to better isolate the trend signal in ice storage decrease over glaciers in Greenland and Antarctica, given by GRACE. (Crowley et al., 2006) also noticed storage trends in the Congo and hypothesised that the basin was drying, although this 'trend' could well have been the result of persistence of dry and wet years, known to occur in this region. The reproduction of these hydrological phenomena prove that the model concept and chosen parameter set yields enough memory in the system to result in interannual variability of the storage process. This is not trivial. Many

hydrological model structures are not capable of yielding interannual variability, especially during low flow periods, because they always return to the same initial condition in the end of the dry season.

The largest difference between LEW and bias corrected GRACE over the targets is attained at the end of the (southern hemisphere) wet season 2004, whereas the differences at the end of the wet season of 2005 and 2006 are significantly smaller. This difference can be attributed to either the LEW model or the GRACE data. For instance, a poor quality of rainfall input (which is satellite based) from the beginning of 2004 can cause significant errors in the LEW model output. This would also partly propagate into the computed bias. Alternatively, it is possible that GRACE did not capture the water storage variation over the target areas very well in spring 2004, due to a poor orbit geometry (Winsemius, Savenije, Van de Giesen, *et al.*, 2006). This would cause an additional bias in GRACE monthly amplitudes, which cannot be corrected for with the GRACE fields used. However, investigation of the ground-track coverage over the basin in March and April 2004 showed that the latter explanation is not very likely. Fig. 5.14 shows this ground-track coverage for the month April 2004 and reveals a quite evenly spaced ground-track coverage in time. March 2004 has a likewise pattern.

The maximum difference between bias-corrected GRACE and LEW can be observed in the Zambezi + Congo area, during the end of the Southern wet season of 2005. At first glance, this is unexpected as the bias is the smallest for this area and the quality of GRACE should improve with increasing size of the target area. The discrepancy should probably be explained by poorer performance of the LEW model, which may have the following reasons:

- It is not entirely convincing that the model concept, encompassing the perceptual model of the upper Zambezi, also fully applies to the Congo. This catchment largely consists of tropical rain forest which may result in very different hydrological behaviour than the more Southern located Kalahari sands.
- Data availability is poor in the Congo basin. There are no discharge time series available in the Congo except for the most Northern tributaries of the river (see Fig. 5.10). Hence, it was virtually impossible to apply a detailed calibration in this region.
- The quality of the FEWS rainfall estimates is strongly dependent on the amount of locally available and operational rain gauge data, assimilated into the rainfall estimate. There are very few of them available in Southern Africa and there are none available in the Congo basin, which means rainfall may be a large source of uncertainty.
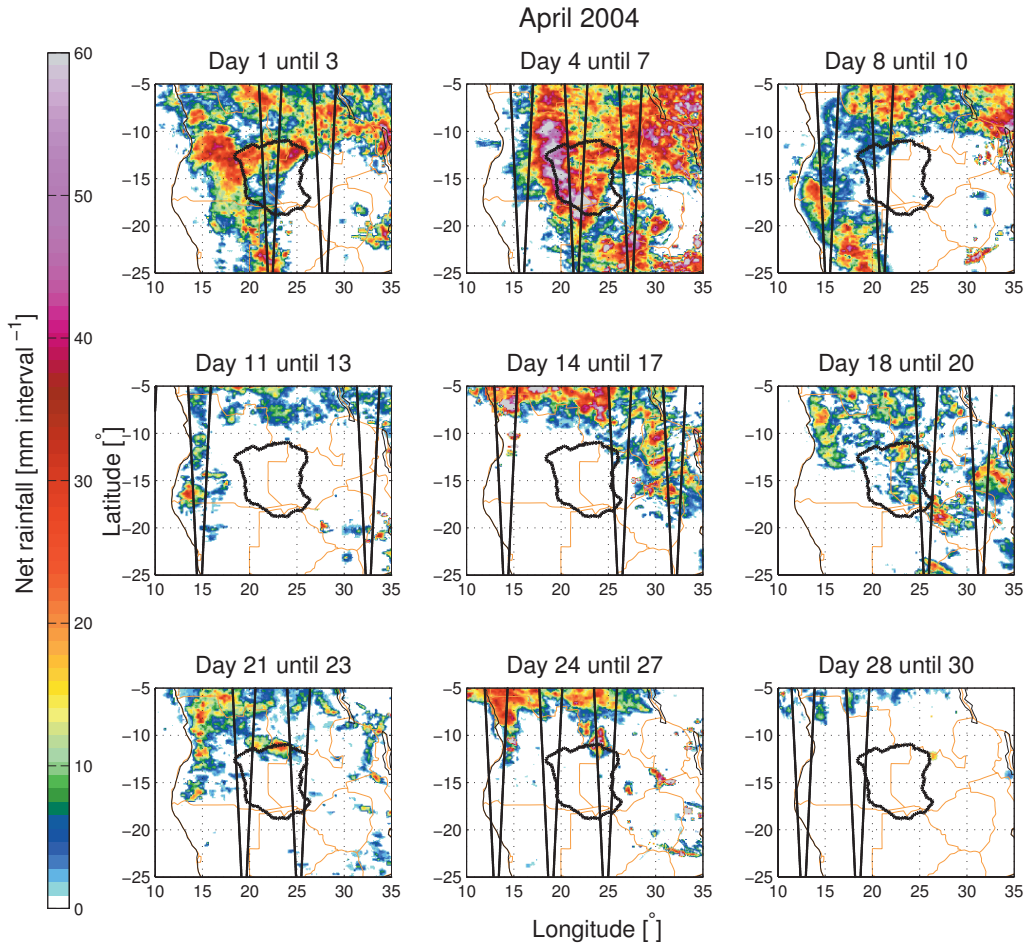
Figure 5.14: Ground track coverage during 9 intervals of 3-4 days in April 2004. The target area is delineated in black. Black dots indicate the ground tracks.

Regarding the regions South of the upper Zambezi, we have much more confidence that the model structure, principally developed for the upper Zambezi, is valid. The Kalahari stretches further South and both the Kalahari and Namib deserts receive very little rainfall and these regions are expected to generate very little discharge. We know from experience and previous modelling attempts (e.g. Hughes *et al.*, 2006; Wolski *et al.*, 2006), that the Okavango river is the only river that has permanent flow in this area. This flow is furthermore known to be released in the Okavango delta, which is therefore modelled explicitly. In terms of storage variability, these characteristics mean that we may expect that all rainfall eventually evaporates. In fact almost all of this evaporation is likely to

take place within one month after the rainfall took place, either from intercepted water or by transpiration from the top soil. This fast transpiration also causes swift blooming periods in the desert, known to occur within a few days after a rainfall event. This means that the storage change in these areas is likely to turn up in the within-month noise of GRACE solutions, rather than in the signal. In short, we may expect that the selected model structure is reasonable in these domains, and furthermore, although calibration was only marginally possible, storage variability is simply easier to simulate (at least on the monthly basis), given that there is far less rainfall here than in the North.

West of the upper Zambezi (i.e. the middle and lower Zambezi) a lot of discharge data are available. This means that refinement of the calibration was possible here, but we have to be aware that at the time of study, no field trip was done in this area and the Kalahari sands do not exist here, which implies a higher uncertainty of the model structure. Nevertheless, also when the whole Zambezi is considered, still a remarkable resemblance between GRACE, corrected for bias, and LEW can be observed.

Finally, some statistics of the annual water storage variations have been computed for the biased and bias-corrected GRACE estimates and compared with the annual water storage from LEW. An annual sine has been fitted through the storage time series and the amplitudes of GRACE, bias-corrected GRACE and LEW have been compared. The results are summarised in Table 5.2. Although interannual variability is not quantified in this way, it is clear that in any case, bias-corrected GRACE estimates of the annual and the monthly mean water storage variations fit significantly better with LEW estimates than uncorrected GRACE estimates. For instance, when using 1000 km Gaussian smoothing, the annual difference reduces from 73 mm to 0 mm for the upper Zambezi, from 50 to 11 for the Zambezi, from 52 to -2 for the upper Zambezi + Okavango and from 38 to -16 for the Zambezi + Congo area, although it is doubtful if a sine fit is an appropriate benchmark for a river basin, where in some areas counter-phase rainfall seasons occur and in some areas even rainfall throughout the whole year. The resulting annual amplitudes are not significantly different for different filters applied. Summarising, for the upper Zambezi and upper Zambezi and Okavango target areas, the relative difference reduces to almost zero, regardless the filter applied, whereas for the Zambezi and Congo, the difference is about 22%. At monthly scale, the different filters do perform differently. A higher correlation length probably results in lower degrees of GRACE noise in the basin-scale solutions, which improves the final results.

Table 5.2: Average annual amplitude of water storage variation. From left to right: GRACE estimate; bias; bias-corrected GRACE; LEW model output; difference between bias corrected GRACE and LEW model output; relative difference between GRACE and LEW model output.

Target	GRACE [mm]	Bias [mm]	GRACE corr. [mm]	LEW [mm]	diff. [mm]	diff [%]
Gauss filter, 1000 km						
UZ	82	73	155	155	0	0
Z	83	61	144	133	11	8
UZO	68	50	118	120	-2	2
ZC	33	22	55	71	-16	22
Gauss filter, 800 km						
UZ	102	55	157	155	2	1
Z	100	46	146	133	13	10
UZO	80	39	119	120	-1	1
ZC	40	15	55	71	-16	22
Gauss filter, 600 km						
UZ	124	36	160	155	5	3
Z	117	30	147	133	14	11
UZO	91	28	119	120	-1	1
ZC	47	8	55	71	-16	22

5.6 Discussion

5.6.1 The information content of field surveys and GRACE

In this chapter, an application of the framework, presented in Section 4.2 has been presented. First, a hydrological model has been constructed, based on prior knowledge, collected during a field survey. The field work clearly helped to develop understanding of the dominant processes, storage compartments and how they are interlinked. Through hydrological reasoning, we came to the conclusion that wetlands are dominant factors in creating connectivity between streams and that they function as a fill-and-spill system. The limited availability of old monthly discharge time series helped to test several model structures, based on the perceptual model that was formulated following the conclusions, drawn from the field visit.

The large spatio-temporal variability in the hydrological cycle results in a great amount of bias due to spatial filtering of the monthly GRACE storage estimates. In fact, this bias has been shown to significantly reduce the information content of GRACE: it not only reduces the average annual amplitude of storage (given that a sine curve may be fitted through an annual storage pattern) but also reduces interannual variability present in the GRACE signal. The reduction of this information content is clearly revealed in

the bias estimates, provided by a prior hydrological model. Given that this bias estimate, relative to the storage, is a function of spatial heterogeneity in estimated water storage, it is important to note that a sufficiently high spatial resolution has to be provided by the prior hydrological information (including the rainfall inputs used) to result in a proper bias estimate. If this resolution is too low (i.e. lower than the spatial extent of the filter applied), the bias may be underestimated. It is important to note that there have been significant developments in the development of better spatial filters. Enhanced spatial filter functions, may prove to reveal a lower bias. Klees *et al.* (2008) for instance, showed that the application of their optimal filter, which is anisotropic and non-symmetric, reveals a bias of only 15% over the Zambezi basin. This is a major improvement and is expected to result in improvements in the field of hydrology because a smaller filter size may be used and the influence of bias on the final hydrological signal is reduced. From a hydrological point of view, a 15% reduction in annual variability may still be significant. If such a bias is disregarded, this may imply making a large relative error in the variability of other terms in the water balance, namely evaporation and discharge, since these may prove to be a small percentage of the total annual rainfall.

This conclusion also has a bearing on the information content that may be expected from GRACE by hydrologists: GRACE is in many cases used to validate the averaged annual storage amplitude. Ironically, many models may be able to make a good estimate of the climatology. But it is in many cases for hydrologists far more crucial to search for other, possibly less straightforwardly quantifiable, signatures of behaviour in GRACE data. Let me refer to the case presented in this chapter, where specific wet and dry seasons exist. Here, it was shown that hydrologically meaningful signatures may also be the interannual variability in the depletion towards the end of the dry season, and the amount of storage increase towards the end of the wet season. This may well provide a validation of the capability of a hydrological model to reproduce such crucial variability and of the system storage thresholds, represented by certain parameters in the model. This is what in fact makes a hydrological model meaningful, especially when prediction of future change is concerned.

5.6.2 GRACE for model structural improvement

Let me now turn to the use of GRACE for model structural enhancement. In the presented case study, we (i.e. the co-authors in Klees *et al.*, 2007) were in fact lucky to end up with an almost unbeatable result, especially for our target area, the upper Zambezi, but even for most of the remainder areas. This success may partly be attributed to the fact that the variability in storage is linked to the occurrence of rainy seasons, and to the fact that we could make sure that discharge was at least reproduced to a certain extent, but what

was not expected is the reproduction of the interannual variability of the storage pattern. It is shown by Winsemius, Savenije, Gerrits, *et al.* (2006), that two model structures, applied on the upper Zambezi (one of them being the LEW model presented here) were in fact able to produce almost the same hydrograph. However, the storage pattern was completely different: not only the amplitudes were mismatched, in addition, the first modelling attempt was not able to reproduce interannual variability in the depletion of storage towards the end of the dry season (i.e. in every year, the storage at the end of the dry season was virtually the same). This would imminently have led to the rejection of this hypothesis as soon as GRACE would have been involved and would have led to the iteration step, described in Section 4.2.

Finally, I cannot guarantee that there are not more model structures that will give a good representation of both ground measurements and orthogonal observations from GRACE. Nonetheless, I would like to stress here that the odds are small that completely different model structures, able to reproduce both signals, will also encompass expert knowledge collected in the field, which one may also consider as a soft form of orthogonal information. It is a pity that in many peer-reviewed journals the editorial process ends up in a struggle for justification of the use of qualitative (i.e. subjective) information in the development of a model structure, although more attempts are nowadays ending up in success (see e.g. Mul, 2009) and the scientific community is starting to accept the fact that 'art' is required in the formulation of better models, especially in ungauged basins (see for instance the opinion paper about 'The art of hydrology' Savenije, 2009).

5.6.3 Recent developments

Of course, a new data source such as GRACE is prone to rapid developments. Throughout the last two chapters, gravity solutions were used that were at the moment of writing state-of-the-art (release 03). Nowadays, GRACE release 04 solutions are readily available (see `http://podaac.jpl.nasa.gov/grace` or `http://isdc.gfz-potsdam.de/grace`). Furthermore, DEOS (TU Delft) develops his own gravity fields including uncertainty estimates (X. Liu, 2008). Post-processing filters for removal of correlated errors are routinely used (Swenson & Wahr, 2006) and far better spatial filters, tailored to the error sources in GRACE have been developed (Klees *et al.*, 2008) which considerably reduce the errors in GRACE. It is therefore likely that repetition of this research would yield even better results. Klees *et al.* (2008) mention for instance that bias due to leakage reduces to about 15% for the Zambezi when the optimal filter is used. As mentioned earlier, this is still a significant amount for hydrology, as important fluxes may represent even less than 15% of annual rainfall. Therefore the proposed method is still a crucial process in the use of GRACE for model development.

Chapter 6

A framework for calibration with hard and soft hydrological information

In the previous chapter, it was shown that GRACE may be used as validation of the model concept, applied in large-scale hydrological modelling. To this end, an iterative framework has been suggested in this thesis, where hydrology and geodesy complement each other: hydrology, by providing prior information about spatio-temporal variability of storage, needed to correct for leakage errors, and geodesy, by consequently validating or rejecting a hydrological model hypothesis and parameter set.

In the following chapters, the focus will shift from the use of satellite information as *validation* and *improvement* of model structures, to its use as complementary information in the *calibration* of hydrological models. At this point, I will abandon the large scale character of GRACE and explore information with more spatial detail. The problems inherent to the use of data in calibration are their uncertainty and the reflectance of this uncertainty in model uncertainty. In my opinion, the uncertainty of most satellite data used in hydrology, is by definition ill-quantifiable and therefore the information derived from it should be treated as 'soft' rather than 'hard'. In other words, satellite data may provide a rough estimate of which uncertainty is quantified by expert judgement instead of concrete numbers with a well-quantifiable objectively determined uncertainty. In fact, in ungauged basins, we have to deal with the fact that *any* available data may be useful, but may also be highly uncertain (without having the means to accurately estimate this uncertainty), intermittent and non-concomitant.

In this chapter, a calibration framework is presented, based on the Generalised Likelihood Uncertainty Estimation (GLUE) that can be used to condition model parameter distributions under these conditions. The framework is capable of handling both hard and soft

information, and can therefore also include satellite information in the calibration process. The essence of the framework is that we use objectively defined limits of acceptability on a number of objectives. This reduces the subjectivity of classical application of GLUE considerably, which otherwise would lead to a cascade of subjective decisions around the defined objectives. Instead of using model residual based information, this chapter shows how to calibrate a given hydrological model by using signatures derived from any available observation. Limits of acceptability of these signatures are derived from inter-annual variability of the signature itself, derived from the data and user perception. First, all hard information with objectively quantified limits of acceptability is introduced in GLUE. This results in an intermediate parameter distribution, which can be used to reduce the sampling limits and to assess which additional information would further constrain model parameters. Soft information may then be introduced, which is here considered as complementary information, being less effective in driving parameter calibration or for which limits of acceptability cannot be objectively derived. Instead, the limits have to be based on the modeller's perception or experience. The framework is demonstrated in an application in the Luangwa catchment in Zambia. Three information signatures are retrieved from an old discharge time series and used to condition model parameter distributions of a conceptual model. Two independent calibration experiments with two significantly different satellite rainfall estimates as model input were performed. In the next chapter, satellite information will be introduced in the calibration framework, as complementary calibration information.

6.1   Introduction

In recent years, a general consensus has developed on the opportunity to profit from any hydrological information that may be available in ungauged conditions. This can be done by effectively combining the information, both to constrain hydrological model parameters and to enhance model structure (Seibert & McDonnell, 2002; Winsemius, Savenije, Gerrits, *et al.*, 2006; Klees *et al.*, 2007; Yadav *et al.*, 2007; Fenicia *et al.*, 2008a). This has also been shown in previous chapters. For information to be used during calibration, multi-objective calibration techniques have been proposed (e.g. Vrugt *et al.*, 2003). However, these can only be applied when dealing with well-defined objective functions. In the presence of significant uncertainty and sparse information, formal calibration methods may be impossible to apply. In fact, a typical problem of many ungauged or scarcely gauged river basins in the world is that even when there is some ground data available, these may be sparse, inaccurate, intermittent, non-concomitant and collected at different time

scales, leading to the fact that it is not clear how one can integrate their non-conventional information content.

In this chapter, a framework is proposed to integrate both hard and soft information delivered by available hydrological observations. The framework makes use of the Generalised Likelihood Uncertainty Estimation (GLUE, Beven & Binley, 1992) within a limit of acceptability approach (Beven, 2006). During the writing of this thesis, an interesting debate was ongoing within the hydrological community about the opportunity of using "formal statistical" versus "informal" methods for model calibration and uncertainty analysis (Mantovan & Todini, 2006; Beven *et al.*, 2007, 2008; Beven, 2008). The latter category of methods includes GLUE which has been often criticised for being subjective. I generally agree with the subjectivity of GLUE but nevertheless GLUE is potentially useful for reducing the uncertainty of hydrological model parameters in conditions of data scarcity, when a formal statistical assessment would not be reliable. In fact, in the well-gauged case, one may a posteriori test the statistics of the error structure, imposed in the definition of a formal likelihood function, if available (input and evaluation) data are in fact continuous and long enough. To strengthen the applicability of GLUE in ungauged situations, a new framework is proposed in this chapter, which fixes the limits of acceptability on 'signatures', which are potentially present in any hydrological data, even when the data is subject to uncertainty, non-concomitance or other problems related to the ungauged nature of the studied river basin.

The use of signatures from time series in model conditioning shows important developments in recent literature. Vogel and Sankarasubramanian (2003) showed that statistical signatures of time series reveal significant and hydrologically meaningful information, which is not conveyed by regularly used least-squares performance criteria. Gupta *et al.* (2008) and Yilmaz *et al.* (2008) proposed to use a set of hydrologically meaningful signatures to evaluate model performance besides less meaningful residual based methods. Yilmaz *et al.* (2008) provide an extensive overview of potentially useful signatures in model calibration. Such signatures have consequently been used in calibration of hydrological models (e.g. Pokhrel *et al.*, in press; Herbst *et al.*, 2009; Bulygina *et al.*, 2009) and model regionalisation (e.g. Yadav *et al.*, 2007; Oudin *et al.*, 2008). Direct calibration or regionalisation may however not be applicable in many basins, because the conventional data sets needed, are not available. Therefore this paper focuses on the use of any information present in non-conventional data sets, available for the basin itself.

6.2 A general framework for integrating hard and soft hydrological information

6.2.1 Problem definition

The focus is on the calibration of a rainfall-runoff model for an ungauged basin, rather than on model structural improvements, such as in the previous chapter. The assumption is that there are at least some observations available to identify one or several objectives for model calibration. Let us say for instance that we have a small amount of discharge observations available. The objectives could be to reach a satisfactory fit of different kinds of behaviour of the hydrograph (e.g. recession curve, peak flows, etc.). Let us assume that the objectives can be specified in the form of target values to be optimised during model calibration. The target values here represent the information content that we retrieve from the data and are possibly affected by significant uncertainty induced by data scarcity or poor data quality. In fact, it is here assumed that the user does not have a conventional data set available, namely concomitant and long series of input and output data observed at the required time step. This implies that a traditional optimisation procedure (either single- or multi-objective), for instance based on least squares of residuals, cannot be carried out.

Nonetheless, the principle goal here is to calibrate hydrological models under the aforementioned ungauged circumstances. Generally speaking, in the ungauged case, formal statistical parameter inference is of limited use since there is not enough data to apply such methods in a rigorous way. In fact, a residual time series is either not available or highly dependent on uncertain input. Either way, a modeller is not able to test the statistical hypotheses underlying any formal likelihood measure.

In this context, GLUE offers promising perspectives to perform model calibration and uncertainty analysis based on any available information that may provide a measure of goodness of fit of model response. This is the basic concept of Bayesian model inference, where subjective probabilities are used to assess the plausibility of model output. A crucial issue is to define the probabilities within GLUE. These should be derived based on demonstrable valid statements. Valid statements are first of all required to be consistent with the probability calculus, of which the axioms were clarified by Kolmogorov in the 1930s (Rougier, 2007). In this chapter, suggestions are made on how evidential probabilities may be defined in ungauged basins, based on the concept of limits of acceptability for the model output (Beven, 2006).
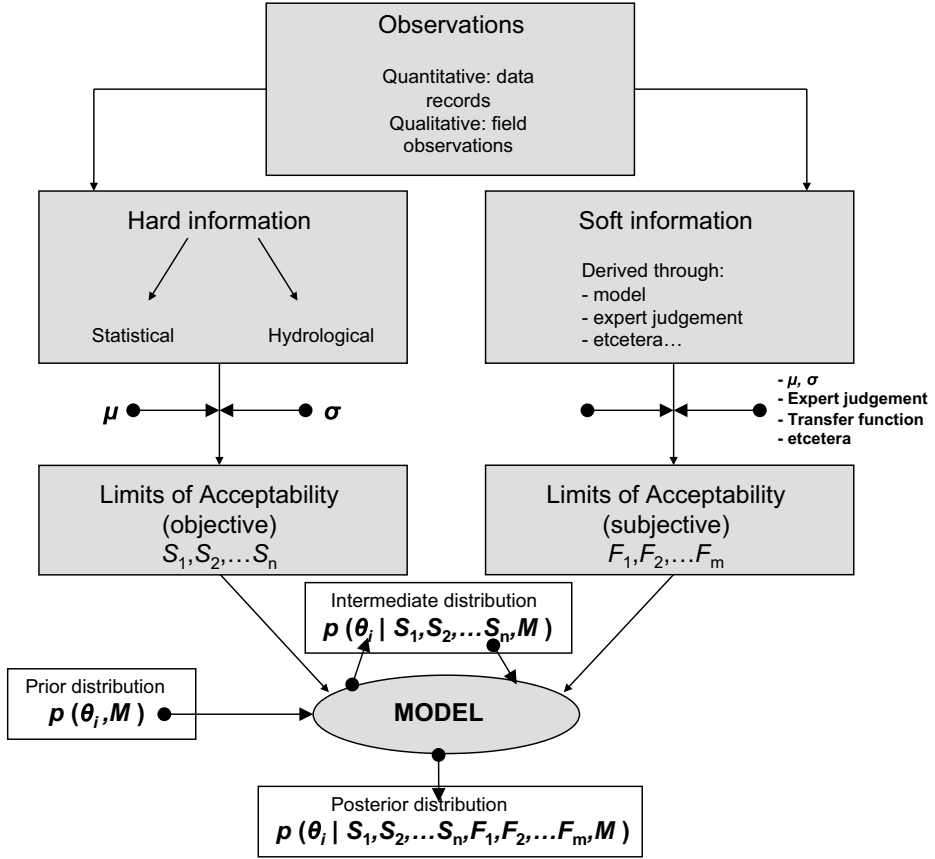
Figure 6.1: A schematic diagram of the proposed framework.

6.2.2 Description of the calibration framework

The proposed calibration framework is based on GLUE and is graphically presented in Fig. 6.1. To reject or accept models, created during Monte Carlo sampling, we have to consider what information may be derived from any available observations related to the rainfall-runoff transformation in the given basin. This information may then be encompassed in the parameter inference process by identifying one or several objectives along with their related target values for the rainfall-runoff model. This means that, during calibration, the model output should satisfactorily reproduce the target in order to be accepted as a plausible realisation of the process. The targets are estimated based on both statistical and hydrological signatures of available river discharge data or any other observations. Many examples of such signatures are given by Yilmaz *et al.* (2008) and Yadav *et al.* (2007). But now the question is: when is the objective 'satisfactorily re-

produced'? In the classical GLUE approach, a number of realisations is computed by drawing random parameter sets from a prior (usually uniform) distribution. Each model that has a lower evaluation score than a certain subjectively chosen threshold, is then rejected and a posterior distribution is estimated from the remaining models and their associated likelihood. The selection of a rejection threshold is rather subjective. Other than in classical cases (Beven & Freer, 2001), I propose to define target values in the form of 'limits of acceptability' of the derived information signatures. The approach to limits of acceptability has been outlined by Beven (2006), but has since then not been applied yet in practice. This approach abandons the definition of a target value (i.e. an informal likelihood function) that summarises all residual information into one mapped variable (such as the root mean square error or Nash and Sutcliffe efficiency). Instead it suggests to evaluate the uncertainty of the observations, used for evaluation, in most cases a discharge time series, and to reject any model that does not reproduce the observed process within the defined uncertainty bounds (i.e. limits of acceptability) of the observations. These uncertainty bounds could for instance be based upon the uncertainty in the rating curve, used to transfer an observed hydrograph of water levels into discharges. From this point onwards, the limits of acceptability approach may be used to find a set of behavioural models. The modeller should henceforth accept each of these models as equally likely. Limits of acceptability therefore provide an alternative to fixing threshold values for informal likelihood measures within GLUE to identify behavioural solutions (Beven & Freer, 2001).

But of course we are dealing with the problem that there simply may not be enough information available about a river basin to be able to construct a concomitant time series of (uncertain) observations and a modelled process, as suggested above. But still, a lot can be done in these conditions. This is the essence of the proposed framework, which is stepwise described below.

1.   First the modeller searches for information content in the form of signatures in any data that is readily available (many examples of signatures are provided by Yadav *et al.*, 2007). The signatures do not necessarily have to be fully independent, meaning that the information conveyed by them could be partially overlapping.

2.   The information is then subdivided into hard and soft information. Hard information is such that a) it has a related target value which has a significant impact on the parameter conditioning of the hydrological model; and (b) limits of acceptability can be objectively defined for the target value itself. This last point is a crucial step as the choice for limits of acceptability may impose a strong control on the outcome. In a later section, I describe how to deal with the subjectivity of this choice. Soft information is considered to be of a complementary nature, which

means that a) the information is less effective to condition the model parameters and/or b) its uncertainty cannot be objectively quantified.

The distinction between hard and soft data may imply some subjectivity which is difficult to completely remove when dealing with ungauged basins. Hard information can be further categorised in statistical properties, such as the autocorrelation of the river flows, and hydrological information, such as the time to peak, the water balance and the shape of a recession curve (Merz & Blöschl, 2008a, 2008b).

3. Monte Carlo simulations are performed while using the limits of acceptability of the hard information to establish whether or not a randomly sampled parameter set is feasible given the information content. Any model that performs within the chosen limits of acceptability is accepted as equally likely, because there usually is not enough information to assign a likelihood to the accepted models. Having sampled a satisfactory number of parameter sets, for each parameter the posterior density may be derived according Bayes' theorem

$$p\left(\theta_i \mid S_1, S_2, \cdots, S_n\right) = \frac{L\left(S_1, S_2, \cdots, S_n \mid \theta_i\right) p\left(\theta_i\right)}{p(S_1, S_2, \cdots, S_n)}, \tag{6.1}$$

where θ_i is a parameter of the given model, $S_1, S_2, \cdots, S_n$ are the given targets, $L\left(S_1, S_2, \cdots, S_n \mid \theta_i\right)$ is the probability of meeting the targets conditioned on the parameter θ_i, $p\left(\theta_i\right)$ is the prior probability for θ_i and $p(S_1, S_2, \cdots, S_n)$ is the probability of meeting the targets. In practice, $L\left(S_1, S_2, \cdots, S_n \mid \theta_i\right)$ is estimated from the empirical density of the accepted parameter sets. $p(S_1, S_2, \cdots, S_n)$ is defined as a normalization constant, so that the posterior distribution $p\left(\theta_i \mid S_1, S_2, \cdots, S_n\right)$ has a cumulative value of unity, such as suggested by Beven and Binley (1992).

4. Based on analysis of the intermediate parameter distributions, a new Monte Carlo simulation is performed, now using the intermediate results as prior. This allows for a computationally efficient subsequent constraining. Now, the soft information with related limits of acceptability is included for further constraining. Remember that objective uncertainty quantification for the soft information could be impossible (see e.g. Franks *et al.*, 1998; Freer *et al.*, 2004; Winsemius *et al.*, 2008).

5. Analysis of the results provides the modeller insights into which part of the parameter space, and, consequently, which outputs are well constrained by means of the included information and what constraints are still lacking. The user may then decide what information has the potential to further condition the parameters and, if deemed feasible, to collect this information.

6. After the collection of new information, the procedure may be repeated to update the parameter distributions with new targets.

One does not necessarily need to perform a two-stage optimisation and therefore the dis-

tinction between hard and soft information is not strictly necessary. In fact, a simultaneous assessment of the achievement of all the hard and soft targets is in principle possible. The advantages of performing a two stage procedure are the following. First of all, if there is a significant difference among the constraining powers of hard and soft targets, the two-stage procedure is computationally more efficient. Due to the hard information the parameter space is first reduced to a smaller region where the soft information, which may be less efficient in constraining the parameters, can be more profitably used. Second, with a two stage procedure the soft information can be used, according to the user needs, for instance to refine the posterior distribution of selected parameters only. Finally, the intermediate parameter distribution can inform the user which parameters are poorly constrained by the hard targets and what information is needed to further condition these parameters. This is essential because any collection and processing of new data is costly and time consuming.

The framework cannot generalise the used signatures as these are strongly dependent on available data, the dominant processes and related time scales and non-linearities of the basin studied and even the users perception, which may be influenced by the objective of the model itself. However, this is also the case in any classical calibration approach, where the modeller decides on the objective function used. Instead, the generality lies in the method to reduce the subjectivity of rejection criteria for models, and the iterations that the modeller can go through after the first results are obtained. The first results can act as a learning tool to discover what additional information is required for further constraining.

6.2.3 Definitions of the limits of acceptability

Now let me turn to the identification of the limits of acceptability of the target values. It is a delicate step of the analysis because it is usually subjective. For the hard information, we propose a procedure for eliminating part of the subjectivity, which can be applied when a multi-year observation is available. The limits of acceptability are based on the analysis of the inter-annual variability of each target value. In this way one can get an objective indication of the uncertainty affecting the evaluation of the given target.

In detail, the identified target value can be computed from the data for a number of hydrological years, therefore obtaining a sample of targets. Then, the sample is transformed to the Gaussian distribution by using the normal quantile transform (NQT, see Kelly & Krzysztofowicz, 1997; Krzysztofowicz & Kelly, 2000). By assuming that the underlying random variable is stationary, a 95% confidence interval is constructed for each transformed target value to define the related limits of acceptability in the Gaussian domain

(i.e. $\mu \pm 1.96\sigma$, where μ and σ are respectively the mean and standard deviation of the transformed sample). Finally, by applying the inverse of the NQT the 95% confidence interval is obtained for each target. When interpolation is not possible, the inverse of the NQT is computed numerically, by linear extrapolation of the tails of the NQT. In this study, the upper and lower 3 points of the NQT have been selected. More details about the computation of the NQT and its inverse are given by Montanari and Brath (2004). We assume that the signature is a stationary process. Of course, this assumption does not hold if significant changes in the catchment have occurred during the observation period.

Given that the limits of acceptability are defined on annual values of the corresponding target values, the objective function should be evaluated for each hydrological year for which the model is run during calibration. Only the parameter sets for which the simulation satisfies all limits of acceptability in each simulated hydrological year should be retained.

The above procedure may seem somewhat crude, especially when only a small number of years of observations is available, but the crucial point here is that a great deal of subjectivity is removed, which especially in ungauged conditions is imperative. In the case that only one target can be identified from the observations used, removal of some subjectivity may not have a great impact, but in ungauged cases, for instance the case described in the remainder of this chapter, it is important because not one, but many objectives are taken from the observations. Subjectively chosen limits of acceptability may therefore give very different results.

6.3   Application to the Luangwa river, Zambia

In order to demonstrate the proposed framework in detail I turn to a real world case study of one of the sub-catchments of the Zambezi, the Luangwa river basin. The intention in this case study is to calibrate a rainfall-runoff model, which should be applicable in estimation of river flows at daily time scales. This could be useful for instance for flood prediction and operation of the downstream reservoir, lake Cahora Bassa. The Luangwa basin (Fig. 6.2) is a relatively pristine and remote area of about 150 000 (km)^2, located in Zambia, Southern Africa. The Northern, most upstream part of the basin is mountainous and is subject to many locally generated flash floods. The downstream parts consist of sandy/loam soils (among which black cotton soils) covered by typical tropical savannah vegetation such as Miombo, Mopane (Frost, 1996) and acacia species. In the lower reaches, dambos can be found that are also used for land cultivation. The North-Eastern boundary (the Muchinga escarpment) is densely forested and interspersed by pristine wetland areas
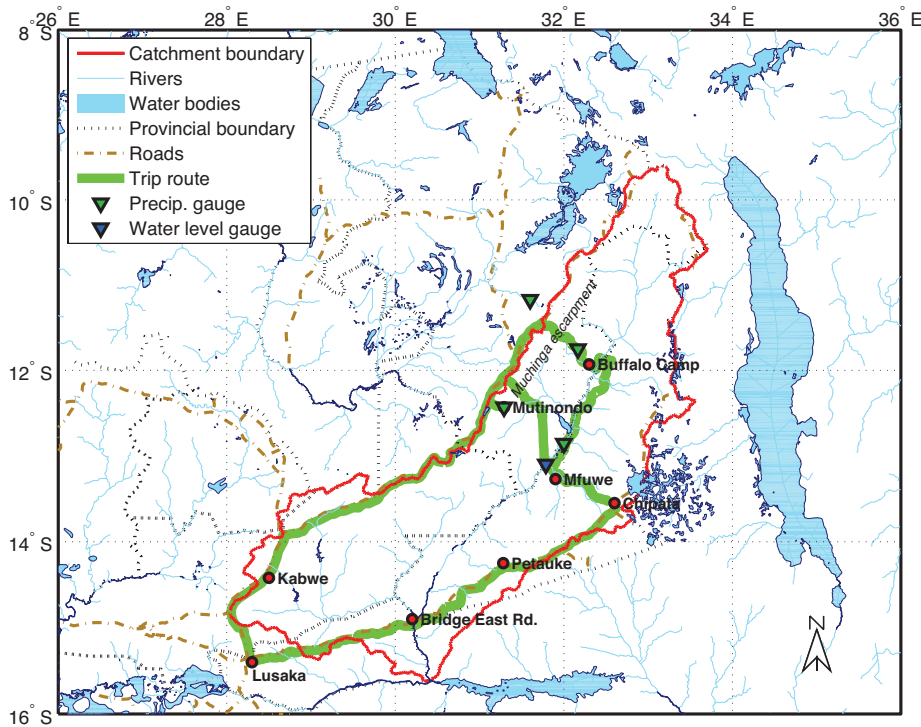
Figure 6.2: Luangwa basin, located in Southern Africa. The study area is plotted in red. The large water bodies and larger dambos are indicated in blue. The tarred road network is also plotted to emphasize the remoteness of the area. All smaller untarred roads inside the basin are only accessible in the dry season. At the bridge on Great East Road, a long old time series of daily discharge data is available.

and large basalt lava rocks. An impression of this pristine area is given in Fig. 6.3. This area has a different hydro-climatology from the low lying savannas. Temperatures on the escarpment are much lower and given the presence of the type of vegetation and dambos, these areas have a higher capability of retaining moisture during the dry season than the lower savannah regions. The average rainfall in the catchment is around 1000 mm yr^{-1}. Rainfall is concentrated in one wet season from November until April.

6.3.1 Data availability

A daily river discharge time series is available near the outlet of the basin at the bridge on the Great East Road (see Fig. 6.2). About 20 years of daily records are therein available from 1956 until 1976. Data were collected by observing the river stage that was subsequently converted to river discharge by means of a rating curve. The rating curve is unfortunately not available. Some of the years are affected by gaps or unreliable data.

Figure 6.3: An impression of the Muchinga escarpment. Photograph, taken during a field visit in July 2008.

A subset of 16 hydrological years was selected (from October to September, 1956–1973, with exception of 1960), with only a minor amount of missing values, that were linearly interpolated.

A considerable amount of monthly rainfall data was found in the GHCN database for the period 1956–1973. Although these records were originally based on daily rainfall, collected by the meteorological department of Zambia in some of the larger towns surrounding the basin, these numbers are no longer available to the public. Therefore, sub-monthly rainfall is not available over the same period as the available discharge records. Instead the two earlier described satellite rainfall estimates at finer time scale for the period 2002–2006 have been considered for running the model: product 3B42 of the Tropical Rainfall Measuring Mission (TRMM, Huffman *et al.*, 2007) and the CPC/Famine Early Warning System (FEWS) daily estimates (Herman *et al.*, 1997). All available data, along with their observation period and observation time intervals are summarised in Table 6.1.

Table 6.1: Summary of the available hydrological information

Variable	Data source	Obs. interval	Start	End
Discharge	Dept. of Water Affairs MWED Zambia*	Daily	1956	1973
Rainfall	GHCN v. 2 (ground data)	Monthly	1956	1973
Rainfall	TRMM 3B42 (satellite-based)	3-hourly	1998	ongoing
Rainfall	FEWS (satellite-based)	Daily	2001	ongoing

6.3.2 Model setup
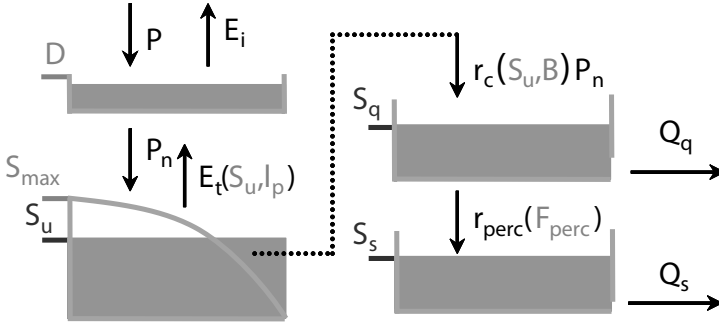
Figure 6.4: Structure of the modified HBV model, adapted from Lindström *et al.* (1997).

A lumped conceptual model has been preliminarily identified for the Luangwa river. It has been derived by modifying the HBV model (Lindström *et al.*, 1997). For the purpose of this study, the structure was slightly simplified into the structure schematised in Fig. 6.4 to obtain a more parsimonious tool in terms of involved parameters. The model structure's storage compartments consist of an interception store, a soil moisture store S_u and two flow generating stores S_q and S_s. Interception E_i is limited by a threshold to net precipitation P_n [L T^{-1}], represented by parameter D [L T^{-1}], by the amount of rainfall P [L T^{-1}] and by the amount of potential evaporation E_p [L T^{-1}]. Therefore, interception can be computed as

$$E_i\left(t \mid \theta\right) = \min\left[P\left(t\right), D, E_p\left(t\right)\right].\qquad(6.2)$$

Then, net rainfall is estimated as

$$P_n\left(t \mid \theta\right) = \max\left[P\left(t\right) - E_i\left(t \mid \theta\right)\right].\qquad(6.3)$$

Throughfall is transferred to an unsaturated soil zone, corresponding to the HBV soil zone, with storage S_u [L], whose outgoing fluxes are controlled by 3 parameters, referred

---

*. Many thanks to Christopher Chileshe, Peter Chola and Adam Hussen for providing these valuable records.

Table 6.2: Uniform prior parameter ranges for the modified HBV model.

Parameters	Description	Unit	min	max
D	Interception threshold	mm day^{-1}	2 (fixed)	2 (fixed)
S_{max}	Unsaturated zone capacity	mm	100	2500
l_p	Moisture stress fraction	-	0.1	1
B	Runoff generation power shape	-	1	5
F_{perc}	Maximum percolation rate	mm day^{-1}	0.01	4
K_q	Recip. of fast reservoir residence time	day^{-1}	0.01	0.5
K_s	Recip. of slow reservoir residence time	day^{-1}	0.001	0.1

to as S_{max} [L], B [-] and l_p [-] (they are equivalent to the abbreviations 'FC', 'BETA' and 'LP', respectively, as used by Lindström et $al.$, 1997). In detail, outgoing fluxes are transpiration E_t [L T^{-1}] and fraction of recharge r_c [-]. The former is computed as

$$E_t (t \mid \theta) = \min \left(\frac{S_u (t \mid \theta)}{S_{max} l_p}, 1 \right) E_{tp} (t), \qquad (6.4)$$

where E_{tp} is the amount of potential evaporation left over after evaporation of intercepted water i.e. $E_p (t) - E_i (t \mid \theta)$. The recharge is determined by the fraction of recharge r_c [-], given by

$$r_c (t \mid \theta) = \left(\frac{S_u (t \mid \theta)}{S_{max}} \right)^B. \qquad (6.5)$$

Recharge $(r_c P_n)$ is transferred to an upper zone with capacity S_q [L]. Stream flow is generated from this zone, assuming that it behaves as a linear reservoir with residence time $1/K_q$ [T]. Q_q represents the fast flow generated from water bodies or seasonal wetlands (dambos). Finally, a lower zone with capacity S_s [L] (conceptualising groundwater) receives a maximum amount of percolation r_{perc} [L T^{-1}] from the upper zone, determined by the parameter F_{perc} [L T^{-1}], according to the relationship

$$r_{perc} (t \mid \theta) = \min (S_q (t), F_{perc}). \qquad (6.6)$$

This zone also empties as a linear reservoir, contributing to the base flow (Q_s) with one parameter $1/K_s$ [T] representing the average residence time. Table 6.2 provides a list of the model parameters along with the respective prior ranges of their uniform distributions.

In order to apply the calibration framework herein proposed, after analysing the available hydrological data, the following objectives have been identified, along with the related target values, to be used to drive parameter estimation:

1. slope of the recessing limb of the hydrograph. Hard hydrological information;
2. spectral properties of non-concomitant daily river flows. Hard statistical information;
3. monthly water balance estimates based on old monthly averaged records of rainfall and stream flow. Soft hydrological information.

The computation of the above target values along with the related limits of acceptability is described below.

6.3.3 Target values and limits of acceptability

Slope of the recession curve

The low flow behaviour during flow recession periods is a property that is insensitive to rainfall forcing. In many cases the recessing limb of the hydrograph can be described by a linear storage-discharge relationship and therefore generally plots as a straight line after a log-transform of the discharge. The offset of this straight line is dependent on the initial storage condition. Lamb and Beven (1997) showed that recession periods may be combined into a master recession curve, which consequently may be used in model calibration, specifically for low flows. Fenicia *et al.* (2006) used this concept in a step-wise calibration of a conceptual model, showing that deviation from the straight line is caused by percolation and capillary rise. Depending on the climatic conditions, one may find fixed recession periods. Particularly in the tropics, seasonally defined dry periods without significant rainfall in each hydrological year can be identified by the modeller. Furthermore, in these periods the discharge contribution has a quite limited range, which mainly depends on the offset of the recession curve within the dry-season period.

In previous applications, recession curve analysis was primarily used to construct a master recession curve without considerations of uncertainties herein. To construct limits of acceptability, a number of sampled slopes can be derived from a number of recession periods. It is hypothesized here that the variability in the slope of the recession curve can be caused by uncertainties during river flow measurements (Di Baldassarre & Montanari, 2008), variability in the spatial distribution of soil moisture in the catchment at the beginning of the recession period or other natural variability, unaccounted for in the model structure. The effect of other causes such as abstractions are expected to be small since the basin is not (yet) heavily exploited by agriculture or industries that consume a great deal of base flow.

Within the present application, given that the climate over the Luangwa river basin is characterised by one dry period per year, the slope of the logarithm of the recession was computed on a yearly basis from the daily river flow data collected in the period 1961

until 1972. based on the NQT of the distribution of yearly sampled slopes and discharge contributions, using the method described in Section 6.2.3. A 95% confidence interval has been used. A parameter set is assumed to be behavioural if it produces a river discharge simulation whose yearly slope of the recession curve falls within the related limits of acceptability.

Spectral properties of the daily river flows

A relevant hydrological signature is provided by the spectral density function of discharge time series. Montanari and Toth (2007) described a maximum likelihood calibration procedure for rainfall-runoff models based on matching the spectral density of the modelled and observed discharge (see also Whittle, 1953). Montanari and Toth (2007) proved that the Whittle likelihood is a powerful measure for model performance under the assumption of stationarity. Moreover, the estimator can be applied when the observed rainfall forcing over the basin is not concomitant with the available river discharge record. The Whittle likelihood (L_w) can be computed by

$$L_w\left(\theta\right) = \exp\left[-\sum_{j=1}^{N/2}\left\{\log\left[J_M(\omega_j,\theta) + f_e(\omega_j,\Phi)\right] + \frac{J(\omega_j)}{J_M(\omega_j,\theta) + f_e(\omega_j,\Phi)}\right\}\right], \qquad (6.7)$$

where J and J_M are the periodograms of the observed and modelled streamflow respectively for the Fourier frequencies ω_j based on N observations, J_M being dependent on the model's parameter vector θ. f_e is the spectral density of the error model, which can be estimated by fitting an autoregressive model on the rainfall-runoff model residuals with parameter vector θ. The interested reader is referred to Montanari and Toth (2007) for more details.

Let us assume that f_e is identically null and that the spectral properties of daily river discharge are largely constrained by a lag-1 autoregressive process. In this case, one may note that the estimator given by Eq. 6.7 reduces to matching the mean value, μ_Q, standard deviation, σ_Q and lag-1 autocorrelation coefficient, $\rho_1(Q)$, of observed and simulated river discharge series (Montanari & Toth, 2007). Therefore, μ_Q, σ_Q and $\rho_1(Q)$ allow us to define a three-elements target vector to resemble the spectral calibration given by Eq. 6.7. Again, the related limits of acceptability can be estimated from the inter-annual variability of the targets (see Section 6.2.3). A parameter set is assumed to be behavioural if it produces a river discharge simulation of which these yearly statistics fall within the related limits of acceptability for each year of the model simulation.

Monthly water balance

Finally, a soft information signature has been considered, that is delivered by an auxiliary rainfall-runoff model at monthly time scale. GHCN provides monthly ground station rainfall data in the period 1956–1973, during which daily river discharges at the basin outlet are also available. GHCN provides monthly rainfall for many parts of the world in historical periods so these data are potentially available in many basins, therefore allowing the computation of monthly mean areal rainfall over many catchments. These rainfall estimates have been used to calibrate an auxiliary rainfall-runoff model running at monthly time scale, to reproduce estimates of mean monthly discharge for any period for which monthly rainfall is available. The daily time step model can then be constrained towards reproducing (in a statistical sense) the long-term discharges provided by the monthly auxiliary model (for more information about the use of auxiliary models, see Seibert, 2001; Schaefli & Gupta, 2007). As auxiliary model, a modification of the 5-parameter rainfall-runoff model HYMOD (Boyle, 2000) has been used and automatically calibrated with the self-adaptive differential evolution algorithm (Brest *et al.*, 2006) on the time period 1956–1973. The first year has been used as a spin-up period.
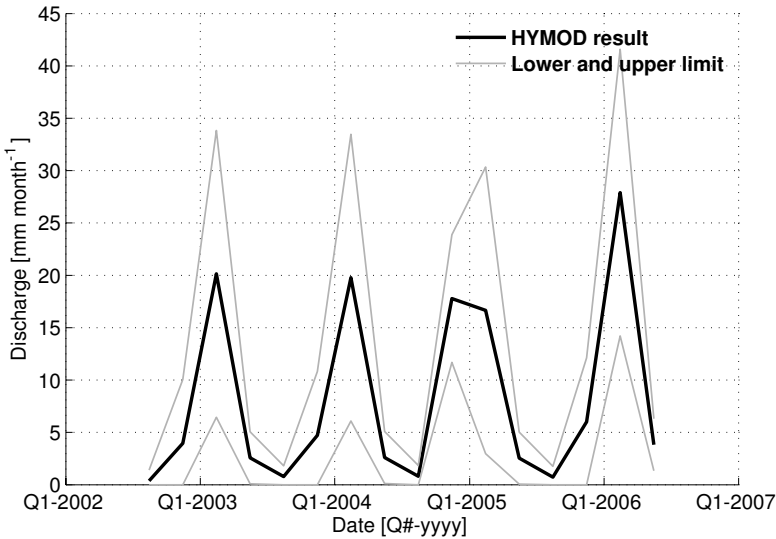
Figure 6.5: Limits of acceptability of seasonal-averaged discharge for the period 2002–2006 based on the HYMOD model, running at monthly time steps, using FEWS satellite rainfall as input.

The model allows for a reconstruction of the monthly discharges at the basin outlet for the period 2002–2006, the run time period of the daily modified HBV model for which daily satellite rainfall observations are available that can be averaged to monthly values.

The estimated monthly discharges were then used to constrain the daily modified HBV model simulation. The reasons why this type of information is considered as soft are the following: first of all the monthly water balance is less effective for parameter constraints compared to the two previous target values; second, the limits of acceptability have to be constructed through an auxiliary model, rather than directly from available observations. Furthermore, the accuracy of the rainfall used for the calibration of the auxiliary model may be strongly related to the density of the rain gauge network. The estimates may not prove accurate enough to warrant hard constraints.

Limits of acceptability for the monthly discharges were computed by dividing the hydrological year into 4 seasons: November–January, February–April, May–June and July–October. For each of these seasons, the season-averaged residuals of the modified HYMOD model were computed, over the calibration period 1956–1973. The inter-annual variability of these residuals allowed us to estimate the related 95% confidence limits around the predicted value for each season. The predicted values are here produced by running the modified HYMOD model with the same rainfall, averaged to monthly values, as used to drive the daily modified HBV model. Fig. 6.5 shows the seasonal limits for the period 2002–2006 and the case that FEWS satellite rainfall is used. Any model realisation of the daily modified HBV model that gives any seasonally averaged discharge estimate outside the limits of acceptability presented here, has therefore been rejected.

The derived limits of acceptability of each target value are given in Table 6.3.

Table 6.3: Limits of acceptability, based on the Normal Quantile Transform. The water balance limits of acceptability are dependent on the output of the monthly HYMOD auxiliary model. Therefore, only the deviation $(+/-)$ from the modeled output is given. Fig. 6.5 shows the time varying limits of acceptability in case FEWS rainfall is used as input.

Description	$\downarrow$ lim	$\uparrow$ lim
Recession slope [day^{-1}]	0.0055	0.014
$\rho_1(Q)$ [-]	0.968	0.994
σ_Q [m^3 s^{-1}]	269	1943
Water balance: Nov–Jan [mm month^{-1}]	-4.24	$+10.69$
Water balance: Feb–Apr [mm month^{-1}]	-20.7	$+12.4$
Water balance: May–Jun [mm month^{-1}]	-1.84	$+3.3$
Water balance: Jul–Oct [mm month^{-1}]	-0.91	$+0.92$

6.3.4 Presentation of the results

Three million parameter sets have been sampled from the prior uniform distributions over the ranges given in Table 6.2. It turned out that D was quite insensitive with respect to the selected targets. Therefore a fixed value was chosen for D. Literature suggests to use a value between 1 and 5 mm day^{-1} for Southern Africa (e.g. De Groen & Savenije, 2006). Pitman (1973) concluded that a value of 1.5 mm day^{-1} should be adequate for many river basins in South Africa. The large size of the basin and full spatial averaging of the rainfall used, suggests that a relatively low value should be sufficient, given the non-linear behaviour of this process. D was fixed on 2 mm day^{-1}. To emphasise the sensitivity of the results with respect to the rainfall input, two calibration experiments have been performed by using the two available rainfall forcings (FEWS and TRMM). Therefore, for each parameter set the modified HBV model was run twice. Model run time was from 1 November 2000 (the start date of availability of the FEWS rainfall data set) until 30 September 2006. The first 23 months of model run time were used as spin-up time. This leaves precisely 4 years for model evaluation (i.e. 1 October 2002 until 30 September 2006). The simulations that satisfied the limits of acceptability for all the considered target values were retained as behavioural.

First, three posterior parameter distributions were derived for each of the two calibration experiments, by considering only one of the objectives in turn. This was done in order to inspect the sensitivity of the model parameters to each objective. The resulting marginal posterior parameter distributions are given in Fig. 6.6 for recession, spectral properties and monthly water balance, respectively. The top side presents the results using FEWS rainfall, while the bottom side presents the case with TRMM rainfall. The figure displays smoothed (for display purposes) histograms over a total of 50 bins, being normalised so that the integral of the density over the 50 bins equals 1.

It is clear that the individual objectives have a limited capacity to constrain the parameter space, but that the combination of the objectives form a powerful constraint. The routing parameters F_{perc}, K_q and K_s are clearly conditioned, all by a different information signature. F_{perc} is well conditioned by the water balance, while K_q is well constrained by the spectral properties. Intuitively, K_s is well conditioned by the recession objective. In fact, K_s is in this model exactly equivalent to the slope of the log-transform of the recession curve, which means that this parameter is automatically cut off at the upper and lower limit of acceptability. The original full prior range ($0.001 < K_s < 0.1$) is therefore not plotted in this sub-figure.

The other three parameters that determine how rainfall is partitioned into evaporation and stream flow, generally exhibit less sensitivity to the objectives used and their poste-
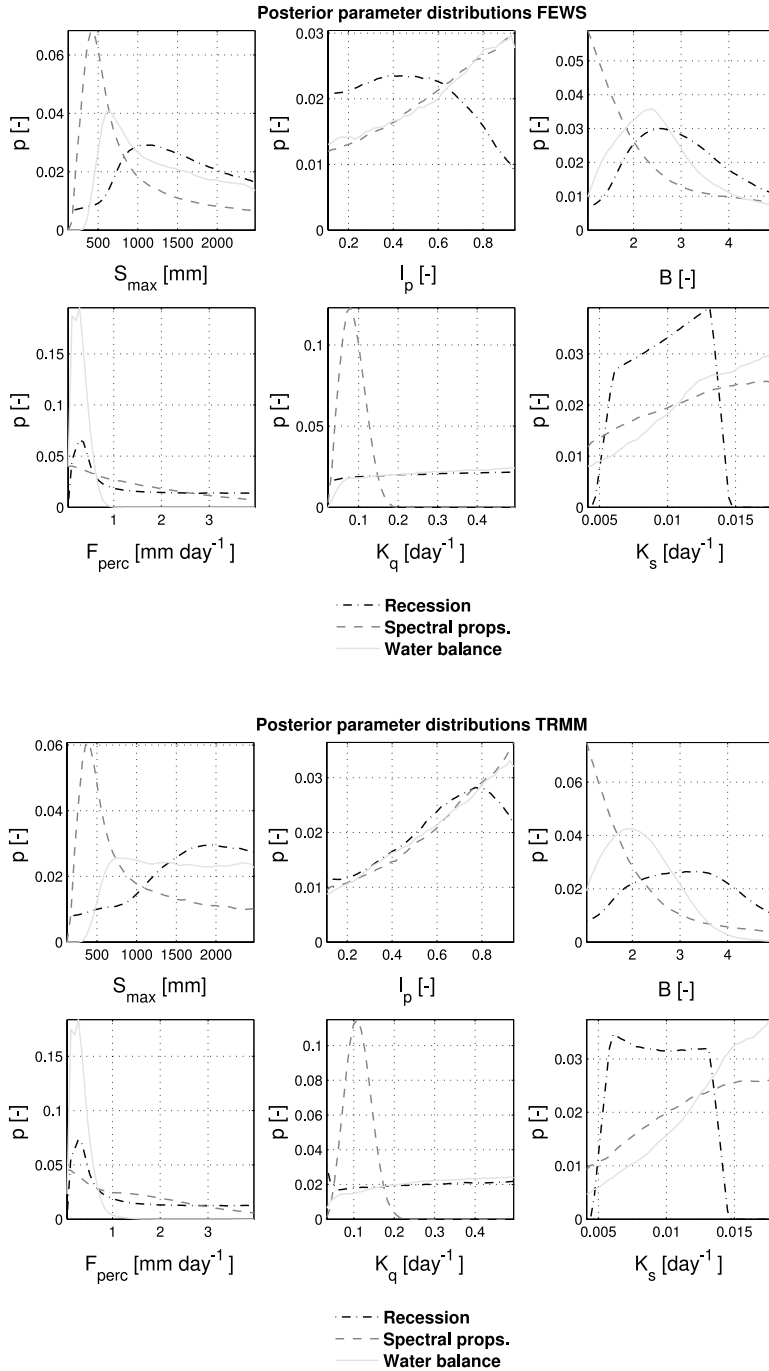
Figure 6.6: Effect of individual targets on the posterior parameter distribution using: (top) FEWS and (bottom) TRMM.
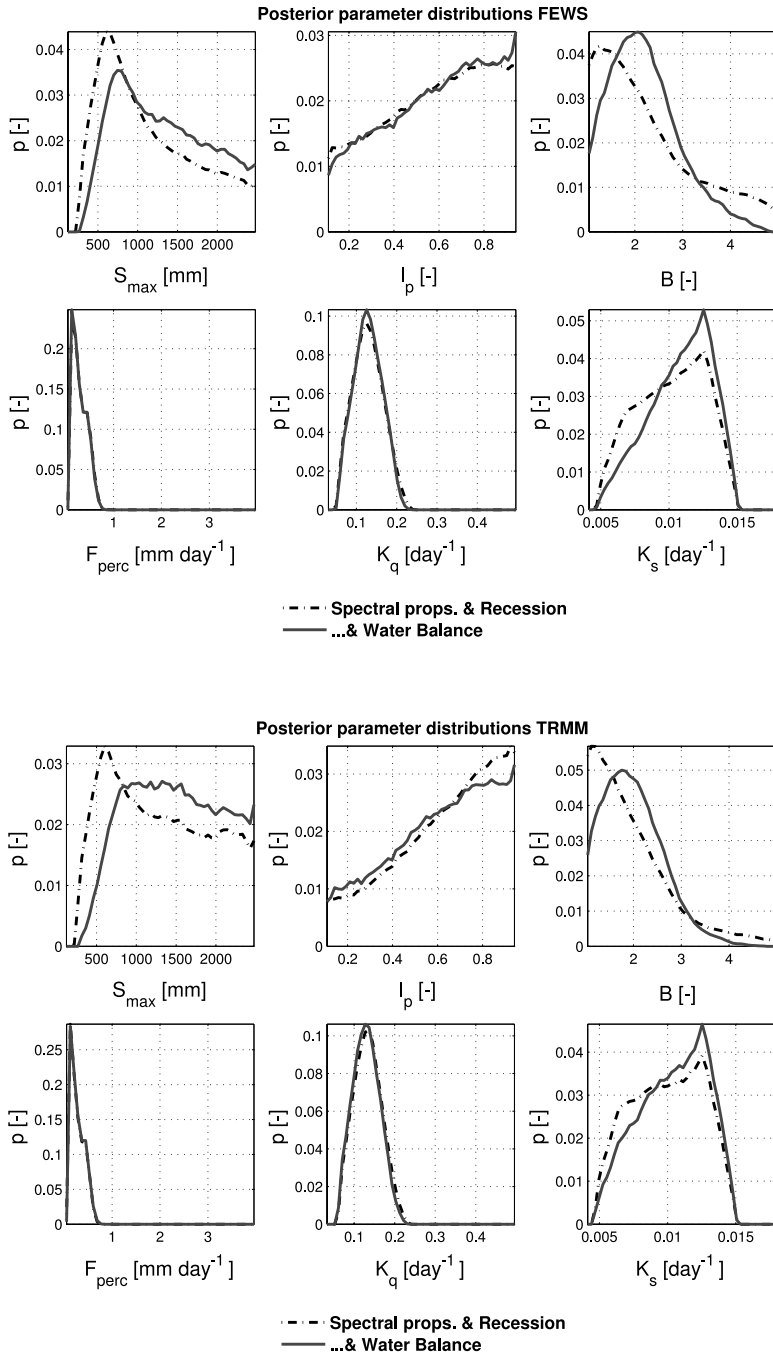
Figure 6.7: The effect of combinations of targets on the posterior parameter distribution using: (top) FEWS and (bottom) TRMM.

rior distributions obtained through the different objectives are somewhat contradicting. Consider for instance S_{max}, which, according to the objective related to spectral properties, shows a much lower mode than the ones obtained by the other objectives. This is not surprising as the different objectives focus on different behaviour of the hydrograph. There are also some slight disagreements between the use of FEWS and TRMM rainfall estimates which may be an effect of their uncertainty.

Fig. 6.7 shows the resulting parameter distributions conditioned on the hard information (black dashed line) and conditioned on all considered information (blue continuous line), obtained by application of Eq. 6.1. Again the top side presents results generated with FEWS rainfall while the bottom side presents results generated with TRMM rainfall. If all objectives are used, about 99.94% of all model realisations had to be rejected but clear constraints could be found for the parameters F_{perc}, K_q and K_s. It is evident that the constraint on the water balance results in a slight shift in the distribution of parameters S_{max} and B. Analysis of the relation between l_p and B reveals correlation (not shown here) which suggests the possible presence of equifinality. It is important to note that the final posterior parameter distributions are quite insensitive to the rainfall estimate used, although the rainfall estimate is significantly different.

The effect of the different constraints on the variability of the outputs becomes apparent in Fig. 6.8. The graphs show what may be referred to as 'behavioural parameter intervals', which reflect the band of the model output encompassed by the behavioural parameter sets. These intervals should not be regarded as confidence intervals because there is still subjectivity involved and only parameter uncertainty is considered (Montanari, 2007). Introduction of the hard data results in considerable reduction of variability in the discharge outputs. Furthermore, it can be seen that the soft information is capable of reducing discharge uncertainty considerably as well and the correlation structure between the parameters changed after introduction of the soft information. The correlation between accepted parameter values of l_p and B increased from 0.19 to 0.65 when soft information was introduced. Finally, it can be seen from Fig. 6.8 (bottom) that the evaporation regime is only marginally constrained by the targets used. It means that the modelled evaporation regime is insensitive to any of the targets considered for model calibration. The only visible constraints can be observed in the dry seasons (July/August/September), when evaporation is strongly moisture constrained. Both the fact that the water balance related parameters are generally poorly identifiable and the fact that evaporation shows almost no sensitivity to the calibration efforts, indicates that additional information should be sought to better constrain these parameters. Given the results, it is likely that information on the soil moisture regime (i.e. conditioning parameters S_{max}, l_p and B, and constraining the evaporation) may provide an additional strong constraint.
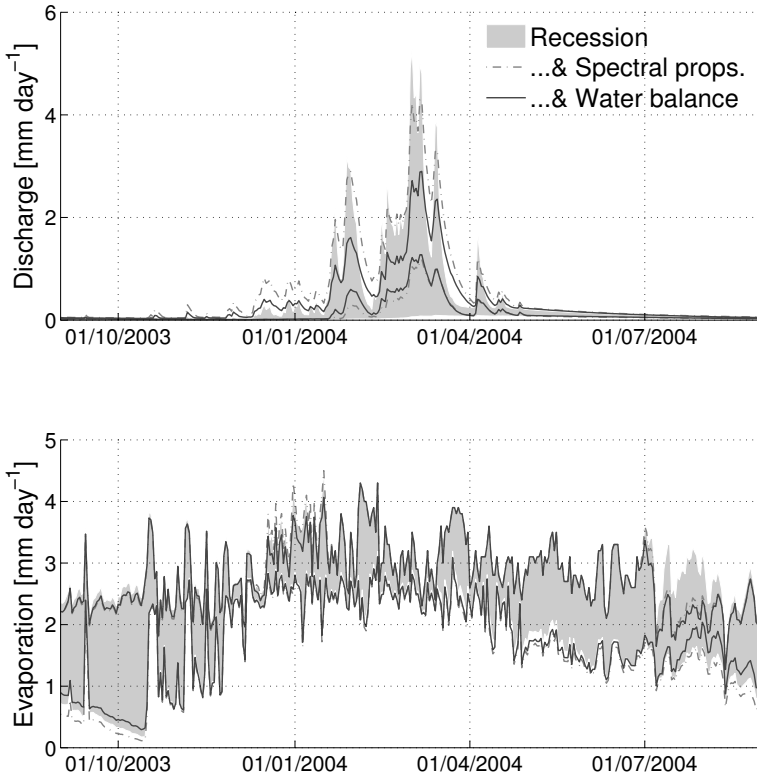
Figure 6.8: Parameter behavioural intervals of the output with multiple constraints: (top) discharge and (bottom) evaporation. The graphs show 5% and 95% behavioural intervals, based on posterior parameter distributions. The graphs show only one hydrological year (2003/2004) for display purposes.

6.4 Validation

For this particular basin, we do not have enough data to perform a meaningful split-sample model validation. In fact, concurrent rainfall and runoff data are not available to check the reliability of the model to reproduce out-of-sample river flow records. This is a common situation in ungauged basins. However, in view of the available information, it is possible to assess the model reliability in reproducing the statistical properties, used for calibration, during a validation period.

In detail, we ran all the accepted parameter sets for a 2-year time span outside the calibration period (October 2006 until September 2008) and computed for each year the hard targets, i.e. the recession slope, recession contribution, lag-1 autocorrelation, mean and standard deviation from the simulated discharge. Furthermore, the auxiliary rainfall-

runoff model has been run for the period 2006–2008 to provide limits of acceptability of the seasonal discharge, as described in Section 6.2.3. Hence, a validation could be made by evaluating how well the accepted models were able to reproduce in a new independent time series the hydrograph behaviours as identified by the limits of acceptability of the signatures. This analysis has been performed with both the FEWS rainfall and TRMM rainfall product. The results are presented in Fig. 6.9 as normalised values with respect
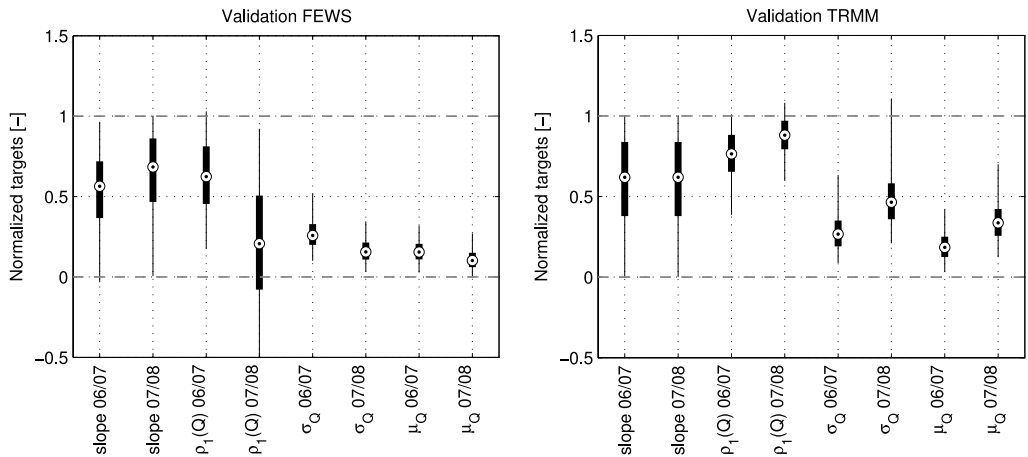
Figure 6.9: Performance of accepted models during validation. The box plots show for each signature and for each simulated year (x-axis, 2006–2007 and 2007–2008) how well the accepted parameter sets remain within limits of acceptability. The box indicates the upper and lower quartile, the whiskers show upper and lower value found in all simulations. For display purposes, the targets have been normalised, so that the lower limits resemble the value 0 and the upper limits 1. The limits of acceptability are shown in red dashed lines.

to the upper and lower limit of acceptability of each target. It demonstrates that most of the parameter sets satisfy all targets in each year. In detail, some inconsistencies were found for the lag-1 autocorrelation in the year 2007–2008 in the case FEWS rainfall is used as input, for which only 68% of the accepted parameter sets remain within limits. Moreover, the limits of acceptability on the seasonal water balance were fully met by 78% of the model realizations when TRMM rainfall was used, and by 97% when FEWS was used. This demonstrates some of the differences, present in the two rainfall estimates and the related modelling uncertainty.

6.5   Discussion

6.5.1 Limitations of the method

A considerable limitation of this framework is that the user needs multi-year information related to each objective to infer the limits of acceptability for the hard information. Moreover, the question remains how reliable the limits of acceptability are when only a limited number of years is available to estimate the variability of the target values. Nonetheless, this method may prove to be useful in many cases. For instance, it is well known that in Africa many hydrometric measurements were in fact taken from colonial times until the mid 70s, which resulted in the availability of quite long historical time series of river flows. Many of them are affected by missing values, or are not associated to concurrent rainfall observations, but nevertheless it was shown that even under these conditions, information can be retrieved from them by means of the framework proposed here. Another possible limitation of the proposed approach is the assumption of stationarity that we introduced in order to compute statistical properties of river flows along a possibly extended time span.

The methodology cannot prevent the occurrence of some subjectivity. For instance, it is up to the modeller what model structure to select, what information signatures to retrieve from the data, and to a certain extent, which information to consider 'hard' and which 'soft'. Subjectivity in model calibration cannot be completely eliminated in the proposed approach. I believe this is unavoidable when modelling ungauged basins. In fact, the need for expert (subjective) knowledge increases with decreasing amount and reliability of the available information. However, I would like to underline that the proposed framework reduces the subjectivity of GLUE.

Although I am aware of these limitations, I feel that the proposed framework allows the user to profit from any available observation to calibrate a hydrological model in a fairly objective manner. This is a valuable opportunity for ungauged basins. In order to profit from any information available a multitude of objectives, individually having a limited capability to constrain a model, is required to provide substantial constraints in the parameter distribution. The control on subjectivity, suggested here, prevents that the large number of objectives used, will result in a cascade of subjective decisions when applying classical GLUE. This would considerably reduce the meaningfulness of any posterior distribution, retrieved from model calibration, which we have prevented here.

6.5.2 What new information should be considered?

The results show that the targets used in the Luangwa case study encapsulate information that allows for constraining of the model parameters. Clearly the constraining abilities of a

target value depend on the hydro-climatological behaviour of the catchment. For instance, in flash-flood dominated catchments, the recession analysis would not have provided a strong control (there may be no base flow at all), while the spectral properties may form a more significant constraint. Therefore, the selection of appropriate model performance measures depends to a large extent on the creativity and perception of the modeller (i.e. hydrology is to a certain extent an 'art', Savenije, 2009). The intermediate and final analysis of the parameter distributions may assist the modeller in the identification of additional useful information that may support the user in planning which additional data to collect. For instance, in the Luangwa River case study the posterior distribution of S_{max}, l_p and B appears poorly constrained while l_p and B are correlated. A significant improvement in constraints would be found, if one of these parameters could be calibrated more independently from the others. This would require a supplementary information source that can provide constraints on other fluxes than discharge. Discharge as a process will always be reproduced by equifinal combinations of S_{max}, l_p and B. Another indication that independent information is required is provided by the bottom panel of Fig. 6.8. The evaporation regime is poorly constrained by any of the information signatures. An orthogonal soft data source that can be considered is GRACE information, although the previous chapters show that its spatial and temporal scale is still too coarse for direct model calibration. Other data, which could provide valuable information for further conditioning are satellite-based evaporation estimates (e.g. Bastiaanssen *et al.*, 1998a). These data can typically provide information on the spatio-temporal variability of the evaporation regime within the catchment, which was shown to be poorly constrained by the chosen discharge-related objectives. Unfortunately uncertainty of satellite-based evaporation estimates in natural catchments is hard to quantify. Therefore they may be considered as soft data (Seibert & McDonnell, 2002; Winsemius *et al.*, 2008) and limits of acceptability should be imposed in a different way than proposed here.

6.5.3 On the value of qualitative observations

In the description of this calibration framework, the suggestion was made to use qualitative observations. Although qualitative information is not used here, it is important to identify the potential use of such information in ungauged conditions. Seibert and McDonnell (2002) show a case study in a well gauged catchment where many qualitative forms of information are used in a fuzzy acceptability framework, because qualitative information is inherently soft. In principle, a modeller can use qualitative information in the calibration in two ways: (a) by superimposing an updated prior parameter distribution; or (b) by superimposing a constraint on the model output. Seibert and McDonnell (2002) use both approaches, dependent on the information considered, in a conceptual model structure

that simulates the physical properties of the catchment well.

In ungauged cases, it is less likely that the model structure is well capable of encompassing all dominant hydrological processes. In such cases, parameters will have to compensate for misconceptions in the model structure. This means that a modeller should be aware that model parameters in ungauged basins may have fewer physical meaning than in gauged basins, which means that it is usually safer to constrain the model output, which will consequently reflect on the parameter inference, than directly constrain the prior parameter distribution, which imposes strong assumptions on the physical representativeness of the parameters themselves.

Let me describe an example that is used frequently in practice: the observation of a certain species of vegetation in a catchment. Many modellers, especially the ones that use a bottom-up modelling approach, try to link such an observation to text-book values of root zone content or root zone moisture capacity and link these to updated prior parameter ranges, without any model inference. The bottom-up modelling approach ignores here two things: (a) certain species may respond very differently under very different climatic conditions (i.e. a textbook value may have been derived in different climatic conditions); and (b) misconceptions in the model structure may lead to parameter values, having to compensate for this misconception as I mentioned above. Ironically, in ungauged basins it is not straightforward to find the right model structure with the right processes. Therefore, as long as there is no evidence that the model structure is truly encompassing the dominant processes, I suggest that qualitative observations should predominantly be used as a constraint on the model output, which will consequently reflect on the parameter inference, rather than directly adapting the prior distribution of parameters. An example: let us say that the observed vegetation species is an evergreen species and we know from literature, that in order to maintain its foliage, the species has to sustain a minimum rate of transpiration per unit of vegetation coverage throughout the year. As a result, we can reject models that do not allow the vegetation to transpire by this minimum rate throughout the year.

6.6   Conclusions

In this chapter, a calibration framework has been proposed which deals with the conditioning of parameter values in ungauged catchments. It can be applied even if rainfall and river flow records are sparse, intermittent, non-concomitant and of poor quality. The framework enables parameter conditioning with highly uncertain data, and non-availability of residual time series. Because formal statistics cannot be meaningfully applied in ungauged

basins, the framework applies GLUE but with a considerable control of the inherent subjectivity of the method, by means of an objective selection of acceptability limits. The framework allows for the integration of hard and soft information in the parameter conditioning process.

In short, it consists of the following steps: a) From available rainfall and discharge observations, hard and soft information signatures are extracted. For hard information, calibration targets, in the form of limits of acceptability, are constructed based on the year-to-year variability of the signature present in the data. Because of the limited number of yearly samples usually available under data scarcity, a Normal Quantile Transform approach is used to construct 95% confidence intervals in the Gaussian domain. For soft information, the limits are either retrieved from a transfer model, if such a model was used to obtain the information, or they are based on the modeler's prior expert knowledge. b) An intermediate marginal parameter distribution is derived for each parameter through Monte Carlo sampling, where models, obeying all the hard targets are accepted as equally likely and all others are assigned a zero probability. c) Further sampling in a reduced parameter space is performed, now including soft information. d) Based on the posterior density of the parameters and the variability in the associated output realizations, the modeler may decide which information is expected to provide an additional strong constraint on the parameter space and if it is feasible to collect this information. e) The modeler includes, after collection, information retrieved from the newly observed variables (not necessarily discharge) in a further constraining step.

In a case study of the Luangwa catchment in Zambia, the use of three information signatures has been demonstrated: the shape of the recession curve, the spectral properties of the discharge process and, as soft information, monthly estimates of the water balance. A calibration experiment with this framework has been conducted, by using two different satellite-based rainfall estimates, both generating considerably different rainfall amounts. While the number of behavioural discharge regimes is seriously reduced by the introduction of the information signatures, the evaporation regime remains marginally constrained. This indicates that additional orthogonal information of the soil moisture or evaporation regime would be needed to further reduce parameter uncertainty. This suggestion is followed in the next chapter, where satellite-based evaporation estimates are employed as additional soft information to further condition model parameters and even allow for distribution of parameters.

Chapter 7

Application of the framework with soft evaporation data

In the final part of this thesis, the application of the framework for calibration, proposed in the previous chapter, is extended with additional soft information. In this chapter, it is demonstrated that remotely sensed evaporation estimates, based on MODIS satellite images and several complementary data, can be used to further condition model parameters. First, evaporation estimates from the dry season in 2006 (April–September) have been used to further constrain those model parameters that are linked to transpiration. Subsequently, the spatial resolution of the evaporation estimates has been used, by allowing spatial distribution of parameters, sensitive to transpiration, based on land cover characteristics.

Finally, a validation has been carried out to test the validity of the proposed framework in combination with all the involved information. To this end, a field mission was organised to collect water level observations from November 2007 until July 2008 at Mfuwe, an internal site in the Luangwa basin. Accurate estimates of rainfall have been derived from a combination of satellite rainfall estimates and rainfall observations collected by local lodge owners and researchers. 100 behavioural parameter sets have been derived from the final parameter distribution and model results have been evaluated against the observations. The validation showed a remarkable correspondence between the modelled and observed discharge, which proves the calibration as valid for the purpose of daily flow estimation. The uncertainty, manifested in the discharge envelopes from the different realisations, encloses the observations most of the time, which also proves that the reduction of subjectivity inherent to the calibration framework, prevented overconditioning.

This chapter is partly based on Winsemius, H. C., Savenije, H. H. G., Bastiaanssen, W. G. M. (2008). Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? *Hydrol. Earth Syst. Sci.*, *12*(4), 1403-1413.

7.1   Introduction

In the previous chapter, it was demonstrated, that any information, even though noisy, incomplete, soft or non-concomitant, may have potential to constrain model parameters and quantify parameter uncertainty. It has been shown that it requires a certain amount of hydrological ingenuity to find appropriate likelihood measures in such cases. For instance, it is not obvious, when a time series of input rainfall is noisy and non-concomitant with other data sources, to seek and utilise other rainfall sources, which may contain less error but are collected at different time scales. Also, not in every case is uncertainty of smaller time scale processes reduced by using the autocorrelation of a signal as an objective. It may depend on the dominant processes and their related time scales and non-linearities, which information content (statistical or hydrological) constrains model parameters most, leading to uncertainty reduction.

It is clear that in ungauged conditions, where most information may turn out to have a weak ability to constrain a model, a large number of objective functions may be needed in the model optimisation process to reach a satisfactory result. In fact, in the previous chapter, I was able to conclude that additional constraints were expected to be found in observations, other than discharge. Note that even under gauged circumstances, this would probably have been concluded, because whether noisy or not, whether concomitant with rainfall or not, discharge has by definition a limited information content, because it is the result of an averaging of processes and, specifically in semi-arid basins, it only represents a small flux in the water balance. Parameter uncertainty (i.e. equifinality) is likely to reveal itself in parameters that by combination reproduce the water balance over longer time scales. To reduce this parameter compensation issue, information should be sought that manifests itself in the inference of parameters that are not only sensitive towards discharge but in addition show sensitivity with respect to the water balance over longer time scales. If the latter parameters can be determined independently, this will reduce equifinality. More practically, in the previous chapter, S_{max}, l_p and B were poorly constrained so it seems logical to look for additional observations, other than discharge, that enable an independent constraining of at least some of these parameters. Since these parameters also have a great influence on discharge, equifinality will decrease as a result.

In this chapter, the results from the previous chapter are improved by including new observations, making use of Eq. 6.1. Additional observations consist of dry-season satellite-based evaporation estimates. The satellite-based evaporation is based on thermal-infrared imagery, which compromises application in the cloud-covered wet season. This is the reason why only dry-season estimates have been used. Previous attempts where evaporation was used for model calibration are presented by Johrar (2002) and Immerzeel and Droogers

(2008). Johrar (2002) calibrated an agro-hydrological model on remotely sensed evaporation rates and validated with in-situ measurements of groundwater levels. Immerzeel and Droogers (2008) calibrated a SWAT model for the Bhima catchment ($\sim$45 000 (km)^2) on remotely sensed evaporation estimates by using a global optimisation algorithm. Here, stream flow is regulated by reservoirs and thus unsuitable for calibration of natural hydrological processes. The advantage of evaporation estimates over remotely sensed soil moisture is that evaporation is a flux, completely equivalent to evaporation from a hydrological model, while soil moisture estimates represent a state, which is not equivalent to the soil moisture state in a hydrological model (i.e. the problem of representativeness of the measurement, e.g. Y. Liu & Gupta, 2007; Wagener & Gupta, 2005).

This study tries to recognise the softness of remotely sensed data (i.e. its ill-quantifiable uncertainty) by using it as a fuzzy, subjective constraint on the parameter calibration framework, described in the previous chapter. The uncertainty of such evaporation estimates may be influenced by many factors, such as uncertainties in the transfer model of radiometric surface temperatures to land surface temperatures; undetected low clouds or aerosols; roughness and emissivity parameterisation and; in the case of the Surface Energy Balance Algorithm for Land (SEBAL, used in this study and described in a later section) the somewhat subjective choice of a "wet" and "dry" pixel as extremes in the surface energy balance. Once this is recognised, dry-season evaporation may from a hydrological point of view, provide the user information in the following form: (a) the contribution of dry-season transpiration to the water balance over a given area; (b) the temporal behaviour of depletion of soil moisture over a given area; and (c) the spatial variability of (a) and (b).

7.2 Surface Energy Balance Algorithm for Land

To derive spatially and temporally variable estimates of the evaporation, the Surface Energy Balance Algorithm for Land (SEBAL) has been applied over the Luangwa river basin. An elaborate description of the current state of SEBAL is given by Allen *et al.* (2007). SEBAL was validated both on field and catchment scale e.g. by Bastiaanssen *et al.* (1998b) and Bastiaanssen *et al.* (2005). It was applied in several studies with varying applications. Some of them are described by Bastiaanssen *et al.* (2002); Schuurmans *et al.* (2003); Mohamed *et al.* (2004); Immerzeel and Droogers (2008); Gragne *et al.* (2008). 15 MODIS TERRA images, ranging from May 2006 until October 2006 (dry season) have been processed under supervision of WaterWatch, Wageningen*. Unfortunately, evapora-

---

*. WaterWatch is a scientific consultancy firm that diagnoses historic and current water management practices at the river basin scale by means of satellite observations

tion cannot be estimated for a complete hydrological year, because during the wet season, no cloud-free images can be found for this region. SEBAL is a residual based energy balance approach in which instantaneous estimates of the energy balance are made based on the energy balance:

$$\rho_w \lambda E = R_n - G - H \tag{7.1}$$

where ρ_w is the density of water [M L^{-3}], λ is the latent heat of vaporisation [L^2 T^{-2}] E is actual evaporation [L T^{-1}], R_n is net radiation, G is ground heat flux and H is sensible heat flux [M T^{-3}]. R_n is estimated from the satellite image derived broadband albedo, Normalised Difference Vegetation Index (NDVI) and surface temperature, together with incoming shortwave radiation estimates from Meteosat Second Generation (see Section 3.3). E, in the dry season, reflects transpiration, soil evaporation and open water evaporation. Open water is hardly present in the dry season and soil evaporation is small compared to transpiration. Hence in the dry season, E is considered to consist of E_t only. G is estimated as a fraction of R_n, being dependent on NDVI. H is determined as an iterative solution to the equation

$$H = \rho_a c_p \left(T_s - T_{a,ref} \right) g_{ah}, \tag{7.2}$$

where ρ_a is the density of air [M L^{-3}], c_p is the specific heat of air [L^2 T^{-2} K^{-1}], T_s is the surface temperature, $T_{a,ref}$ is the air temperature at a reference level [K] and g_{ah} is the aerodynamic conductance to heat transfer [L T^{-1}]. Both g_{ah} and H are a function of wind shear and are therefore iteratively determined. Two known anchor points need to be selected where $H = 0$ and $H = R_n - G$ are fulfilled (i.e. the "wet" and "dry" extremes in the satellite image). Then, the following assumption is made:

$$T_s - T_{a,ref} = aT_s + b, \tag{7.3}$$

where a and b are calibration coefficients that can be found through calibration on the two anchor points. With this linear equation, $T_s - T_{a,ref}$ is found for the whole satellite image. Finally, $\rho_w \lambda E$ is found as the residual of Eq. 7.1. 24-hour evaporation is found by assuming that the evaporative fraction, $\rho_w \lambda E / (R_n - G)$ is constant over the day as given below (with time (t) in hours):

$$\rho_w \lambda \int_{t=0}^{t=24} \frac{E\left(t\right)}{R_n\left(t\right) - G\left(t\right)} \mathrm{d}t = \frac{\rho_w \lambda E_{inst}}{R_{n,inst} - G_{inst}}. \tag{7.4}$$

Here, the subscript "inst" stands for "instantaneous". To yield period-averaged evaporation, daily surface conductance g_s [L T^{-1}] estimates were derived by inverting the Penman-Monteith equation (Monteith, 1981; Penman, 1948):

$$g_s = \left(\frac{s\left(R_n - G\right) + \rho_a c_p \left(e_s - e_a\right)\gamma}{\gamma \rho_w \lambda E} - \frac{s}{\gamma} - 1 \right)^{-1} g_{ah}. \tag{7.5}$$

s and γ are the slope of the vapour pressure curve at given air temperature and the psychrometric constant [M L^{-1} T^{-2} K^{-1}], e_s and e_a are the saturation vapour pressure at given temperature and the actual vapour pressure [M L^{-1} T^{-2}]. All meteorological input required to solve this equation has been taken from downscaled ECMWF fields. A physically based approach has been followed to downscale the coarse grids of ECMWF to near-surface 1×1 (km)^2 variables (see Section 3.4.1). A Jarvis model (Jarvis, 1976) has been used to correct g_s for temporal variability of the meteorological conditions within the period of time (Bastiaanssen *et al.*, 1994; Farah, 2001). This estimate of g_s has been inserted in Penman-Monteith with the concurrent period-averaged meteorological conditions to yield period-averaged (typically 10–15 days) estimates of E over the dry season of 2006, in the remainder referred to as E_o ('o' for 'observed').

7.3   Evaporation as additional soft constraint

7.3.1 Introduction of evaporation in the lumped model

The first experiment, described in this chapter, directly follows the results from the previous chapter, by considering evaporation as a soft additional information source. It is considered soft because

- The evaporation estimates have been retrieved through a model
- Their uncertainty cannot be quantified straightforwardly

Evaporation estimates have been introduced in Eq. 6.1 as a lumped approximation of the evaporation over the entire basin as follows. Two new targets (F_1 and F_2) were introduced, describing the total contribution of evaporation to the water balance, and the evolution of transpiration over time (i.e. target a) and b), mentioned in Section 7.1). The first target is a limit of acceptability, defined as

$$F_1\left(\theta\right) \to 0.9 \left[\int_{t=0}^{n} E_o(t)\mathrm{d}t\right] < \int_{t=0}^{n} E_s(t)\mathrm{d}t < 1.1 \left[\int_{t=0}^{n} E_o(t)\mathrm{d}t\right]. \qquad (7.6)$$

where E_o is the SEBAL evaporation and E_s is the simulated evaporation. Thus, modelled evaporation integrated over time, should be within $\pm\,10\%$ of the integral over time of observed evaporation. The estimate of 10% is based on expert judgement and many studies, performed in the past by WaterWatch. The second target should be evaluated for each time period, for which a satellite image has been processed and is described as

$$F_2\left(\theta\right) \to 0.7E_o(t) < E_s(t) < 1.3E_o(t), \qquad (7.7)$$

Table 7.1: Intermediate extreme parameter ranges, based on the results of Chapter 6.

Parameters	Description	Unit	min	max
D	Interception threshold	mm day^{-1}	2 (fixed)	2 (fixed)
S_{max}	Unsaturated zone capacity	mm	100	2500
l_p	Moisture stress fraction	-	0.1	1
B	Runoff generation power shape	-	1	5
F_{perc}	Maximum percolation rate	mm day^{-1}	0.01	2
K_q	Recip. of fast reservoir residence time	day^{-1}	0.0053	0.0148
K_s	Recip. of slow reservoir residence time	day^{-1}	0.017	0.25

meaning that for each observation period, the modelled evaporation should be within $\pm$ 30% of the observation. The choice for 30% error seems somewhat pessimistic at a first glance, but especially over highly heterogeneous areas, both in terms of land cover and elevation, uncertainties may be high and in cases where quantification remains difficult, it is safer to be pessimistic than to be overoptimistic about them. Also note that, assuming that effective uncertainties in any observation are statistically independent in time, the error should reduce by $\sqrt{n}$ when averaging over n observations in time. If we apply this rule given a random maximum error of 30% per satellite image, the average error is smaller than 10%, meaning that F_1 and F_2 are consistent with each other.

In the previous chapter, the prior parameter ranges (see Table 6.2) could be updated to intermediate parameter distributions. Therefore, sampling is performed more efficiently, if it is performed from this intermediate distribution. The extreme values from the intermediate distribution were therefore selected as new sampling ranges and are given in Table 7.1. In this experiment, again 3 million parameter sets have been sampled. Furthermore, the (in)sensitivity of the introduced information with respect to the chosen rainfall estimate has again been emphasised by performing a calibration attempt with both FEWS and TRMM rainfall as input.

7.3.2 Discussion of the results

Fig. 7.1 shows the intermediate parameter distribution, the distribution, derived when only evaporation would have been available (i.e. introducing F_1 and F_2 in Eq. 6.1), and the full posterior parameter distribution. After introduction of evaporation, S_{max} has a clear optimal value, which is in fact quite similar with different rainfall inputs. l_p and B are not well identifiable, but a correlogram, displayed in Fig. 7.2, reveals that after rejection of models that do not obey the evaporation regime, only l_p and B (i.e. the ones that

were unidentifiable before) have a very clear dependency. This dependency is higher after the introduction of evaporation constraints (the correlation coefficient going from 0.42 to 0.84) meaning that equifinality is now indeed concentrated in the interdependency of only these two parameters, while in the intermediate distribution, more parameters (probably S_{max}) were conspiring in equifinality. In other words, the compensation between the water balance determining parameters S_{max}, l_p and B, and hence the equifinality in general, has significantly decreased. The remainder of the parameters F_{perc}, K_q and K_s are insensitive to the evaporation data.

An important observation is the difference in parameter distributions of l_p and B obtained with FEWS and TRMM. With TRMM rainfall, the hard targets and soft water balance target used in Chapter 6, appear to agree with the evaporation data, where parameter B is concerned. Integrated over time, FEWS produced about 5 to 10% more rainfall in the rainy season of 2005–2006 than TRMM. Let us assume that in the long run, FEWS produced too much rainfall in the rainy season 2005–2006. To reach a wetness state at the end of the wet season that would allow the model to reproduce the subsequent dry-season evaporation within limits of acceptability, B has to have a low value. This is because a low value for B results in higher runoff fractions, which are necessary to reach a satisfactory wetness state. However, B can physically not be smaller than 1. This may be the explanation for l_p, having a high optimal value, which physically imposes a high potential for plant stress. I.e. the relatively large excess water in the soil is not used by the vegetation cover, but instead may remain in the soil. Due to the inconsistency between parameter inference in the two data sources, I wish to hypothesize that FEWS is in fact the inferior rainfall estimate.
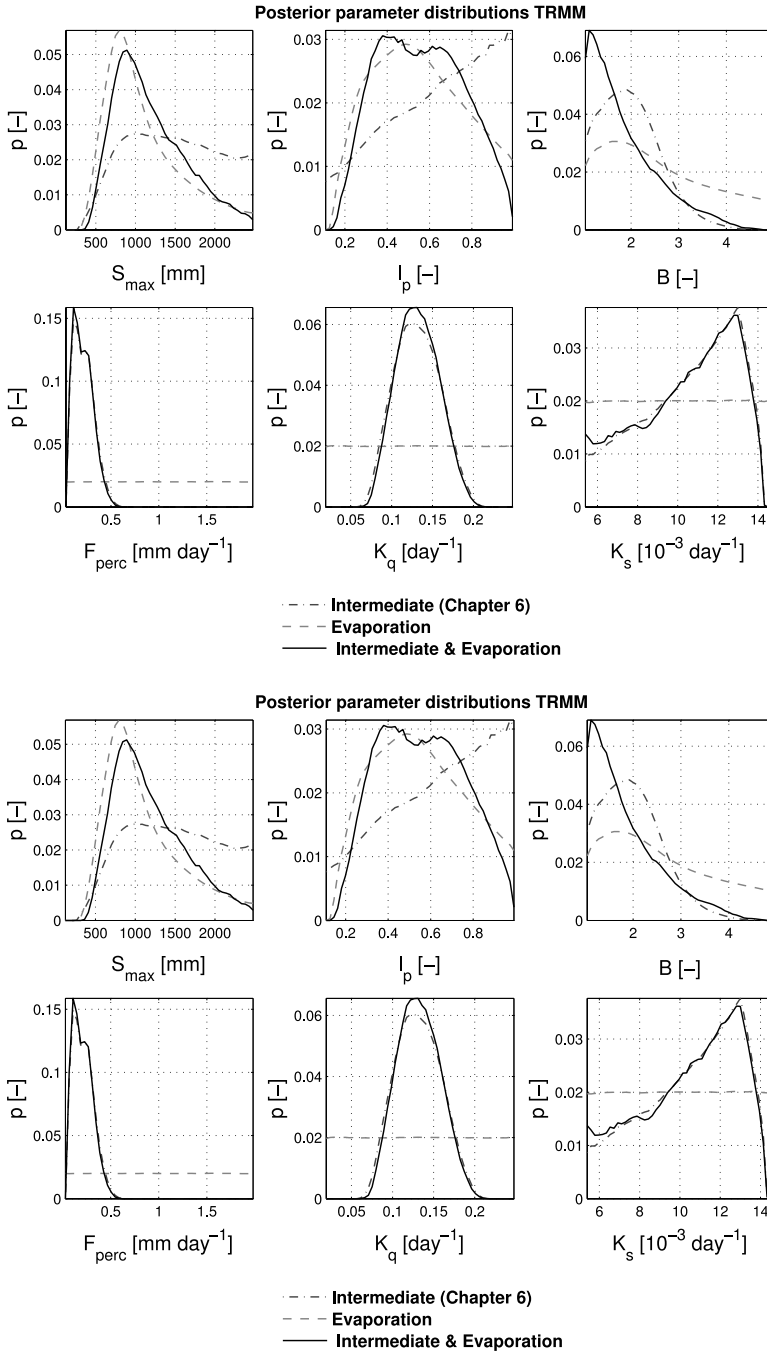
Figure 7.1: Intermediate (i.e. from the previous chapter) and posterior parameter distribution with evaporation-related objectives using: (top) FEWS and (bottom) TRMM.
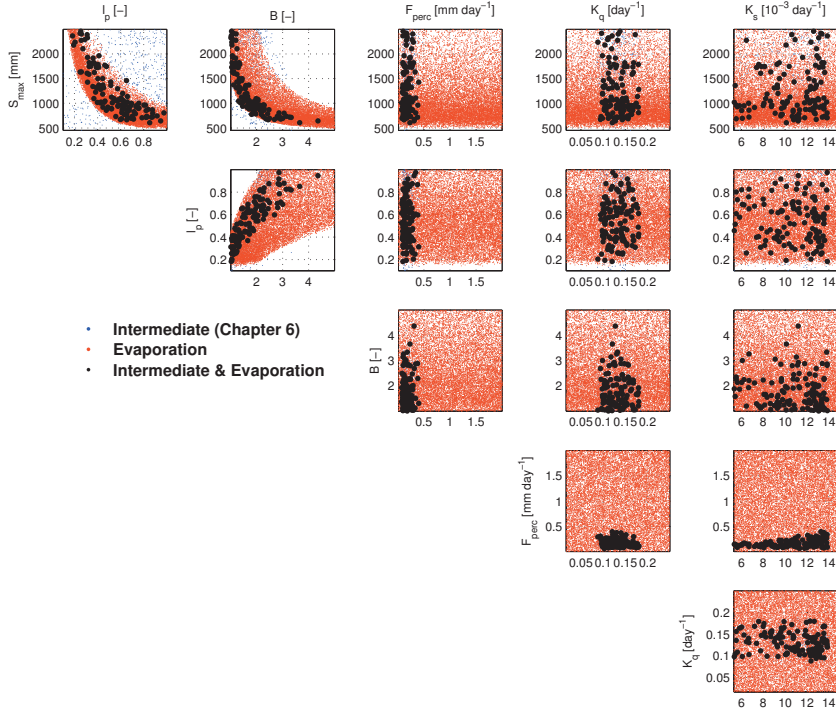
Figure 7.2: Correlogram of parameters from parameter distributions based upon the intermediate results, the evaporation only, and all constraints together.

What remains to be analysed is the effect of the reduction in equifinality on the outputs. Fig. 7.3 shows the reduction in possible realisations of the final outputs compared to the intermediate results at the 90% behavioural interval (see for a definition of behavioural intervals, Section 6.3.4). It is somewhat trivial that the evaporation regime is seriously constrained after introduction of evaporation data. The impact of this additional constraint however, does not manifest itself in further constraints on the discharge realisations when a model is run under the boundary conditions for which it has been calibrated. Discharge remains a small flux in the water balance, which means that small uncertainties in the water balance, may give large uncertainties in discharge. As shown in the top-panel, the plausibility of the discharge regimes appears unaltered. Nonetheless, when climate (i.e. the boundary conditions) changes, the spread in possible discharge realisations is likely to be considerably smaller when constraints in the evaporation regime have been taken into account.
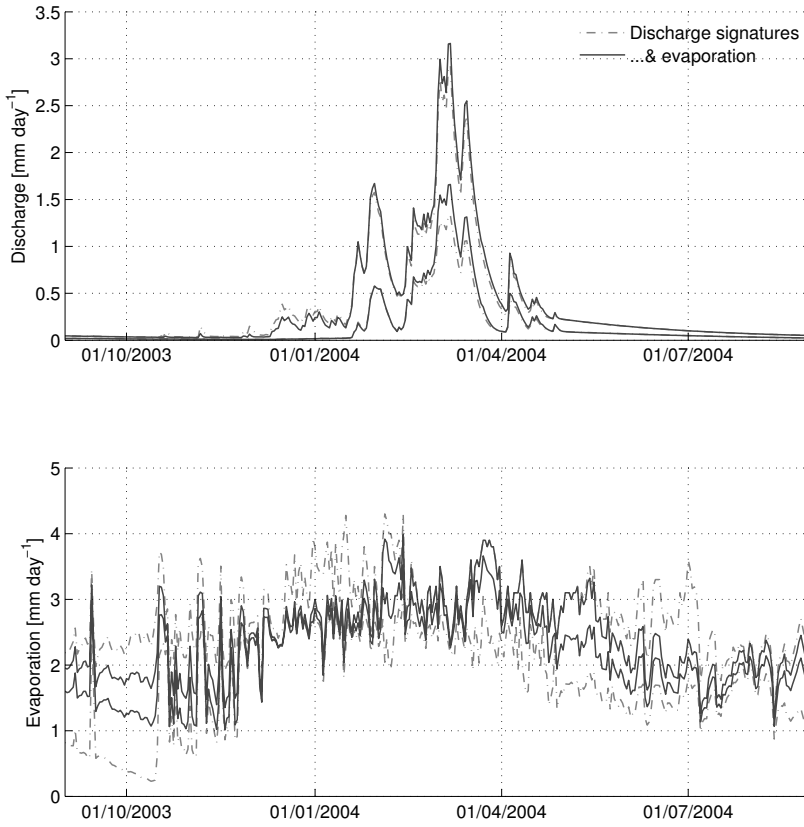
Figure 7.3: Behaviour intervals of discharge and evaporation given only the constraints of the previous chapter, and given the evaporation constraints.

7.4   Spatial disaggregation of transpiration related parameters

Naturally, it is attractive to try to profit from the spatial distribution of data, even if this is at the cost of introducing some equifinality. This may be interesting for many applications, especially the ones that are dealing with internal variability in a river basin (see Grayson & Blöschl, 2001, for an overview of the use of spatial information in hydrological models). Therefore, another experiment has been conducted where parameters that are related to the transpiration regime have been disaggregated into spatially variable parameters. In detail, parameters S_{max} and l_p were allowed to vary in space as a function of land cover, while the remainder of the parameters was considered lumped. A semi-distributed model setup was made. Current state of the art land cover maps do not sufficiently describe the typical land cover of Eastern Zambia. Therefore the spatial distribution of the model

was based on a one-year time series of decadal SPOT NDVI images. An unsupervised classification was performed to differentiate between different land covers. The characteristics of the main dominant land covers was roughly defined through field investigations (the observed routes are displayed in Fig. 6.2), elevation differences and, where available through investigation of high resolution, Google Earth[†] overflight information. Based on this information, four dominant land cover regimes were defined: *riverine*, *dambos* (a local word for wetlands), *forests* and *highlands* as described in Section 6.3. The regions were manually delineated into polygon-shaped model units to decrease the computation time (Fig. 7.4).
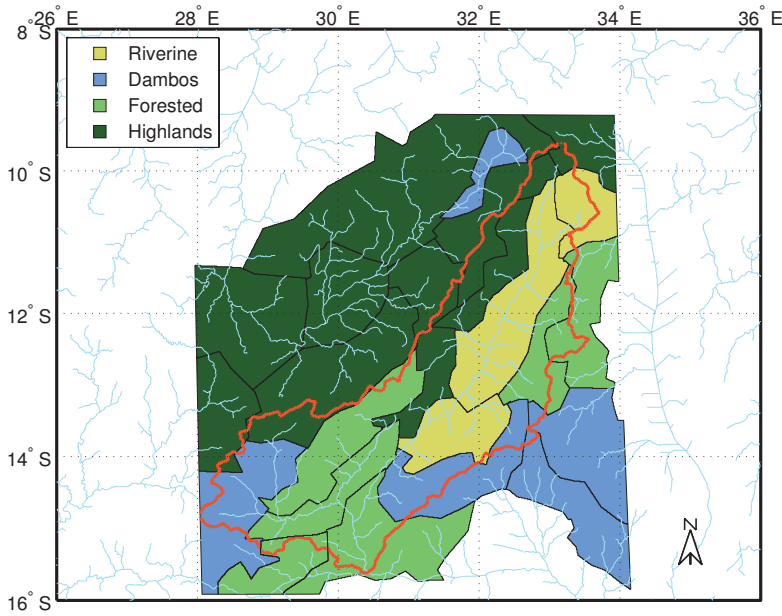
Figure 7.4: The disaggregated model units with their dominant land cover characteristics.

To evaluate S_{max} and l_p for each model unit independently, SEBAL evaporation estimates were lumped per model unit. The available SEBAL evaporation maps correspond to the average evaporation over a period of 10 to 15 days. When more than 25% of the pixels of a model unit within an evaporation map appeared to be cloud-covered (and thus were

---

† http://earth.google.com/

excluded from the computation of the energy balance), the value was assumed to be too uncertain and was thus discarded for the evaluation. The daily evaporation, estimated by the model in each model unit, was averaged over the same periods as the SEBAL estimates. Subsequently a Monte-Carlo framework was applied to estimate posterior parameter distributions for each model unit, given the model structure and given that the quality of the SEBAL data is reasonable and unbiased. Uniform prior distributions were imposed for both S_{max} (varying between 200 and 2000 mm) and l_p (between 0.1 and 1). A model run time from September 2004 (driest moment in the year) until October 2006 was used. This period covers two rainy seasons, which allowed a spin-up of the soil zone for the model evaluation period, May 2006 until October 2006.

By evaluating on smaller target areas, the degree of noise, both in the data used to force the hydrological model, and in the evaporation data, becomes more dominant. This noise may be different in each model unit and may depend on: its size; the local quality of the rainfall inputs; and, considering evaporation, the heterogeneity and unobserved phenomena such as warm clouds or aerosols in the satellite image used. Therefore it is difficult to come up with reasonable and coherent limits of acceptability for the modelled evaporation regime, because these limits may be different for each model unit. It was therefore decided to first perform a Monte-Carlo experiment for each model unit separately, selecting a constant n number of best performing models for each model unit, given the standard least squares criterion,

$$L(\theta_i) = \frac{\sum\limits_{t_p=1}^{m} \left[E_o\left(t_p\right) - E_s\left(t_p\right)\right]^2}{\sum\limits_{t_p=1}^{m} \left[E_o\left(t_p\right) - \bar{E}_o\right]^2}, \tag{7.8}$$

where t_p is a time period [T] for which a SEBAL map was developed, m is the number of observation time periods available and E_o and E_s are the SEBAL-observed and simulated evaporation respectively [L T^{-1}]. L becomes lower as the model performs better. Hence, for each model unit an equal amount of models could be selected as performing adequately, regardless of the amount of residual noise between the output evaporation and the SEBAL evaporation. Given that some model units have the same dominant land cover, the distribution of all model parameter sets of model units with the same dominant land cover could be assessed, and a feasible search space for S_{max} and l_p per land cover for further basin-scale calibration could be defined.

7.4.1 Results

Of each model unit for which evaporation estimates of satisfactory quality were found (some model units were almost permanently clouded), the best 5% of the parameter
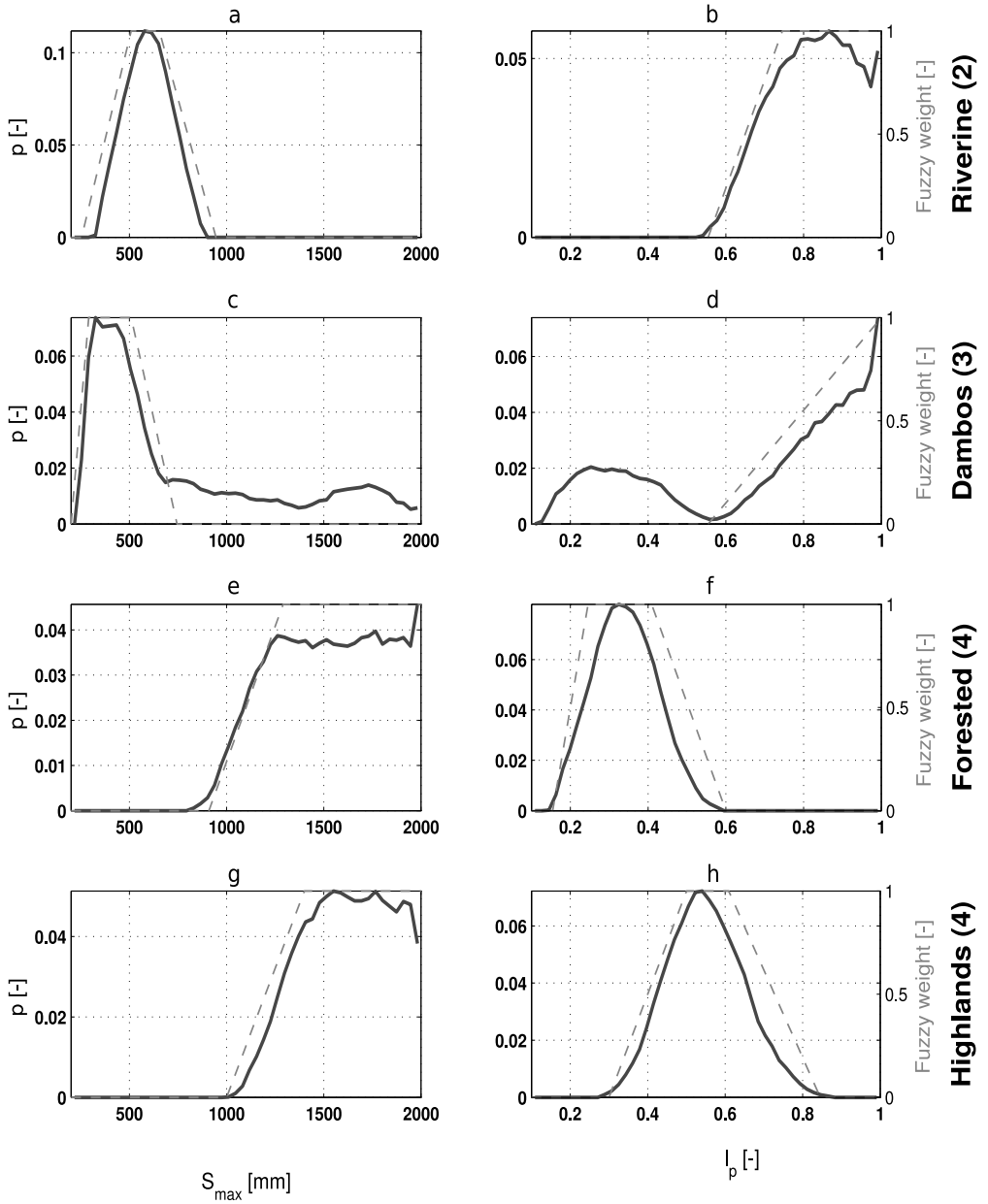
Figure 7.5: Smoothed histograms (50 bins) of all behavioural parameter sets of the transpiration related parameters S_{max} and l_p over each land cover. The bracketed numbers along the y-axes on the right side indicate how many model units were available to construct the histogram for the considered land cover.

realisations was retrieved. An estimate of the resulting parameter distributions of the evaporation sensitive parameters S_{max} and l_p is given in the form of smoothed histograms of the resulting parameter samples in Fig. 7.5. S_{max} has a clear optimal range for all land cover classes. High optimal ranges are found in forested and highland regions, while riverine areas and dambos show relatively low values for S_{max}. Forested model units show a relatively low identifiability, although the parameter ranges suggest that S_{max} should at least be higher than 1000 mm. For land cover classes dambos and highlands, the parameter response surface reveals a clear clustering of optimal parameter ranges for both parameters. Forested and riverine areas show less clustering for l_p. An interesting phenomenon is the apparent bi-modal distribution of l_p in the dambos. Not surprisingly, the second, less pronounced mode is caused by the optimal solutions of one particular model unit, which was presently marked as being dambo dominated. A closer field inspection revealed that this area consists largely of agricultural fields, predominantly cropped with tomatoes, corn and other in Zambia commonly grown crops. This model unit should therefore be modelled with an agricultural model which can take irrigation supplements, crop growth and harvesting into account. This is outside the scope of this study and it is therefore not considered in the remainder of this thesis.

Although the parameters represent area-averaged and thus *effective* values, the relation of the parameter response surface with the identified land cover follows the knowledge we have on the land cover regime. For instance, the dambo dominated areas do not have deep rooting vegetation, since water levels are too shallow in the wet season for deep rooting trees to survive. Shallow rooting grasses, reeds and shrubs dominate these areas and it is well known that the grass wilts very soon after the rains have stopped. This explains the rather persistent optimal value of l_p for all 3 dambo model units being close to 1 (Fig. 7.5d). It implies that transpiration rates decline immediately when $S_u < S_{max}$. A relatively low optimal value for S_{max} (Fig. 7.5c) is found which also concurs with the small root depths of the dambo-type vegetation. Riverine areas show more or less a similar pattern (Figs. 7.5a and b), although the amount of model units used to sample from the parameter space was limited to only two and the parameter clustering is less pronounced. This is clearly visible in Fig. 7.5b, where it is clear that l_p still has a wide optimal range. An obvious reason for this is that not all heterogeneity in soil type and vegetation coverage within the chosen dominant land cover classes is accounted for.

The highlands on the other hand, are covered with (partly evergreen) forests that apparently have a large reservoir of water available for transpiration. l_p is rather low (Fig. 7.5h), indicating drought resistance, and S_{max} (Fig. 7.5g) is high in these areas (2000 mm or even beyond the prior range). There is no optimum found for S_{max} and I believe that a greater prior parameter range would not yield any better results given that the annual

rainfall is in general much less than the maximum prior value for S_{max} of 2000 mm. This means that as S_{max} becomes higher, modelled evaporation becomes less sensitive for variability in S_{max}. The highlands are located on the rather isolated Muchinga escarpment, where the evergreen forests on the edge of the escarpment may act as a sink to which groundwater converges, perhaps even from outside the Luangwa catchment itself. It is therefore likely that the model structure applied, may not be suitable for these areas: first, the *Miombo* woodlands may tap from groundwater, and second, there may be a lateral influx of groundwater, which cannot be modelled with a 1-dimensional box-type model such as the one we present here. A perception could be that the model should be replaced by a 2-box configuration where soil moisture is replenished in the dry season by uptake from the groundwater and discharge is generated from a groundwater reservoir. There are however no measurements done in this area to support this hypothesis.

The posterior parameter distribution for forested regions (Figs. 7.5e and f), shows a quite wide distribution in the combined response surface of the four forested model units. This may be due to the two main forest types in the basin. In the field, we have observed large areas covered with multiple species of *Miombo* (*Brachystegia*), some species suited to the hotter lower areas, some to the colder, higher elevated areas, exhibiting far less seasonality (Fuller, 1999). Some of the lower lying and hot areas are covered by *Colophospermum Mopane*, known to intersperse dambos (Chidumayo, 2005). *Mopane* is known to dominate areas with relatively shallow and poor soils that are not well drained (Lewis, 1991) whereas in areas with more favourable growing conditions, other (for instance *Miombo*) species will dominate. *Miombo* is known to root deeply and use deeper soil moisture or groundwater reserves. The reason for their dry-season dormancy may well be temperature related rather than soil moisture related (Chidumayo, 2005), which could mean that in these woodlands, we should include temperature as a transpiration constraint, which may lead to further model structure improvement. If for instance average temperature will rise in the future, these trees may keep on transpiring at the cost of deeper groundwater. The misconceptualised model presented here, would instead stop transpiring far sooner in this future climate due to moisture constraints. It would require a far more detailed land cover map to identify what type of forest we are dealing with and what its coverage is.

7.5   Validation

The final part of this chapter shows a validation of all the results obtained in the previous chapters by combining all the constraints found in the search for *behavioural* parameter sets of the semi-distributed model. Validation of the final results, presented above would

in gauged circumstances be done by running a random set of models, drawn from the calibrated parameter distribution, with a rainfall dataset and consequently evaluate the results with an available discharge time series. The ungauged nature of the Luangwa river basin does not allow for such a simple validation of these calibration efforts. Recent rainfall observations are scarce, satellite rainfall observations are generally uncertain and discharge measurements in periods with daily rainfall are not available. In this section, I will show how this problem has been overcome.

Figure 7.6: The stilling well at Luangwa bridge, Mfuwe, after rehabilitation.

7.5.1 Available data

It was decided to try to gather a short validation time series of both rainfall and discharges. Ideally, a great number of daily measurements from rain gauges has to be available, concomitant with discharges, taken for instance at Great East Road bridge, where the original discharge time series, used in Chapter 6 was measured. Unfortunately, most areas in this particular river basin are extremely remote and furthermore it was logistically difficult to position a water level measurement device at Great East Road bridge: first of all, this bridge is positioned far away from any collaborative party, and second, no supervision is possible. The latter is important, because in many hydrological studies, gauging instruments are frequently expropriated or vandalised (Kongo *et al.*, 2007; Mul, 2009). Therefore, it was decided to position a pressure transducer with loggers for water and atmospheric pressure at the Luangwa bridge in Mfuwe, bordering the South Luangwa National Park (see Fig. 6.2). In former decades, a stilling well used to be present at this bridge, consisting of a steel pipe, attached to one of the bridge's supporting columns. During one of the Luangwa's ferocious floods, this stilling well was seriously displaced. In November 2007, the stilling well was repositioned[‡] and measurement equipment replaced. Fig. 7.6 shows a photo of the rehabilitated gauging station. Our collaboration with the Zambia Wildlife Authority and a local fisherman[§] ensured that there was plenty of social supervision at the site. Water level measurements were taken from November 2007 until July 2008, capturing one full wet season including parts of the recession period. Fig. 7.7 shows a photo of the successful retrieval of the instruments in July 2008.

An additional problem in this region is the reliability of the rating curve. Unfortunately, the author of this thesis did not have measured rating points available and no rating mission could be planned. Rating in the Luangwa is particularly dangerous during floods, due to the instability of the banks and floating debris. Furthermore, the Luangwa is subject to a large erosive power. This means that especially the rating of low water levels is likely to change rapidly over the years due to changes in the river bed's geometry. The only dataset at hand to reconstruct a rating curve was an old series of water levels with many gaps and the discharges, derived by a rating curve. These measurements were taken from a staff gauge which is positioned on the banks, which may be a dangerous task during flood events. This could have caused the many gaps. In the past, discharges have been derived from these water levels through an unknown rating curve. The Q–h relation

---

Figure 7.7: Successful retrieval of measurement equipment in July 2008.

showed that apparently a rating curve of the form

$$Q = a(h - h_0)^b \tag{7.9}$$

was originally used with Q being the discharge [L T^{-3}], h the water level [L], h_0 a reference no-flow water level and a [L^{3-b} T^{-1}] and b [-] being rating constants of which physically one can show that b should be around 1.7, given that the channel is more or less rectangular in shape. Unfortunately it turned out that from the rating points, apparently a value

for b of about 2.85 was selected by previous hydrometrists which is physically impossible. Therefore, a re-fit was done between the old set of measured water levels and the converted discharges, assuming that b was equal to 1.85, which is a reasonable upper limit for this parameter. h_0 and a were fitted. h_0 was found to be equal to 1.3 m and a equal to 70 $m^{1.15}$ s^{-1}. The relation found was used to convert the presently measured water levels into discharges, after a correction for h_0 for the offset between the reference level of the old measurement instrument and the new pressure logger.

Recent daily rainfall measurements over the Luangwa are only readily available from satellite sources, but their quality is not adequate for a proper validation. Therefore, support was asked from local lodge operators in Luangwa South, on the Muchinga escarpment and Luangwa North and from the Frankfurt Zoological Society (FZS) [¶]. They provided a data set of rainfall from a total of six rain gauges at different sites (3 of them located in the Mutinondo Wilderness area), all located upstream of Mfuwe, that allowed for a considerable reduction of uncertainty in the satellite rainfall estimates. The location of rain gauges is indicated in Fig. 6.2. The uncertainty is assumed to be partly due to bias, and partly due to a spatially correlated error. Essentially, the rainfall correction applied here follows the assumptions:

- Monthly accumulated rainfall estimates from stations are representative for the grid cell monthly accumulated average rainfall rate. This assumption implies that: a) at monthly temporal scale and at the spatial scale of a grid cell, there is no co-variance between rainfall and physical variables such as land cover and topography; and b) over longer periods, randomness of rainfall rate in space reduces (theoretically by a factor of the square root of the number of observed days, if spatial distribution of rainfall events within a grid cell are completely random).
- Ground measurements are unbiased in time. Any error in these measurements is of a random nature.

First, an estimate of systematic bias over the whole target area is computed. Given the assumptions presented above, the maximum possible bias may be estimated as a scaling factor (η_{max} [-]), which can be computed by the following equation

$$\eta_{max} = \frac{\sum_{i=1}^{n} P_{g,i}}{\sum_{i=1}^{n} P_{s,i}}, \tag{7.10}$$

where P_s is monthly space-based rainfall in a certain grid cell, P_g is monthly ground

rainfall [L T^{-1}] and i [-] represents some location in the grid where there is also a ground observation available. However, there are no means to discern systematic and random error straightforwardly, which is why in this thesis, some constraints have been imposed on this scaling factor. The first constraint is based on the coefficient of variation ($\sigma_\epsilon/\mu_\epsilon$) of the residuals ($\epsilon$) between P_s and P_g and is computed as

$$f_1 = 1 - \left| \min\left(\frac{\sigma_\epsilon}{\mu_\epsilon}, 1 \right) \right|. \tag{7.11}$$

f_1 ensures that if the variance of the residuals is small compared to their mean, the likelihood of η_{max} to be indeed a systematic bias over larger areas increases. The second constraint is based on the amount of available stations to estimate the systematic error and is formulated as

$$f_2 = 1 - \exp\left[-\frac{n-1}{c} \right], \tag{7.12}$$

with n being the number of station values available and c a scaling constant [-]. f_2 imposes that as the number of available stations for estimating systematic error increases, the likelihood of η_{max} increases. Here, a value of c equal to 5 has been assumed. The final scaling factor (η) is therefore a reduction of η_{max} by means of f_1 and f_2 in the following form.

$$\eta = 1 + f_1 f_2 (\eta_{max} - 1), \tag{7.13}$$

and the gridded bias corrected satellite rainfall estimate becomes

$$P_{s_1} = \eta P_s. \tag{7.14}$$

The next correction step includes the randomness of discrepancies between station observations and satellite estimates after bias correction. This has been done, roughly following the method presented by Reynolds (1988). This method has been applied before with spatial rainfall estimates from satellites and other sources, e.g. by Xie and Arkin (1996). They apply a merging algorithm based on preservation of the second derivative (i.e. the spatial pattern) of rainfall in space on a monthly basis, in this study provided by the satellite estimate. This can be written as

$$\frac{\partial^2 P_{s_1}}{\partial x^2} + \frac{\partial^2 P_{s_1}}{\partial y^2} = \frac{\partial^2 P_m}{\partial x^2} + \frac{\partial^2 P_m}{\partial y^2}, \tag{7.15}$$

where P_s is monthly space-based rainfall and P_m is the satellite–ground merged estimate. Xie and Arkin (1995) show that on a 2.5° grid, the accuracy of monthly precipitation estimates, compared to a grid cell average should be about 10% when 5 rain gauges are present in the grid cell. In the application presented here, there is only one station
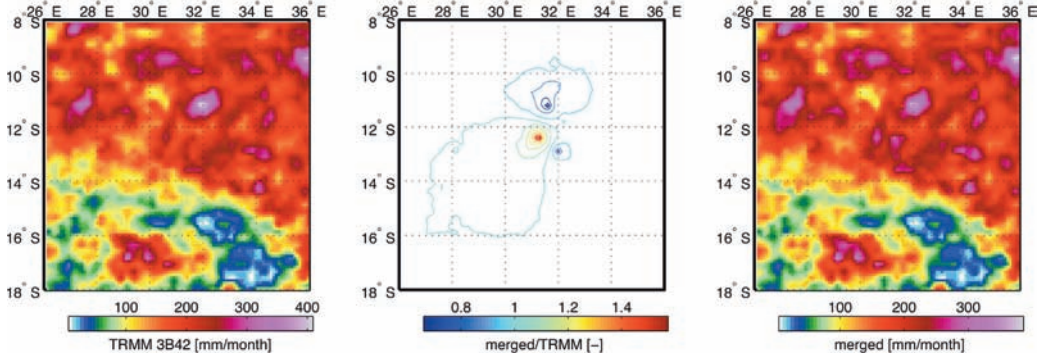
Figure 7.8: Example of the effect of merging monthly rainfall ground records with TRMM rainfall estimates in February 2008, in the middle of the rainy season. From left to right: the original TRMM grid, the relative difference (i.e. merged divided by original) and the merged estimate.

present per grid cell, but the cell size may be up to a hundred times smaller. Therefore it is assumed that this method is valid in this case. Eq. 7.15 may be discretised into

$$\frac{P_{s_1}^{x-1,y} - 2P_{s_1}^{x,y} + P_{s_1}^{x+1,y}}{\Delta x^2} + \frac{P_{s_1}^{x,y-1} - 2P_{s_1}^{x,y} + P_{s_1}^{x,y+1}}{\Delta y^2} =$$
$$= \frac{P_m^{x-1,y} - 2P_m^{x,y} + P_m^{x+1,y}}{\Delta x^2} + \frac{P_m^{x,y-1} - 2P_m^{x,y} + P_m^{x,y+1}}{\Delta y^2}. \tag{7.16}$$

As long as there is no prior knowledge about P_m, this equation is trivial for each grid cell, assuming that the values in the boundaries of the grid are essentially the same as in the uncorrected grid. The boundary conditions of Eq. 7.15 are extended by including knowledge about P_m in the solution at any grid cell where a monthly station value is available. The supply of this information, conditions the grid towards the values, measured at the ground and should provide an optimal monthly gridded rainfall estimate, given that ground truth rainfall is accurate and is, at monthly aggregated scales, representative for the grid cell under consideration. The ratio between the uncorrected and the final corrected monthly grid may be applied on the daily TRMM estimates to provide daily corrected rainfall estimates. An example of the effect of this merging technique is given in Fig. 7.8 where rainfall records have been merged in February 2008 over the Luangwa and its surroundings.

Table 7.2: Evaluation criteria, used to select parameter sets during validation.

Type of data used	limit of acceptability	$\downarrow$ lim	$\uparrow$ lim
Discharge Great East Road bridge	Recession slope [day^{-1}]	0.0055	0.014
	$\rho_1(Q)$ [-]	0.968	0.994
	σ_Q	269	1934
	Water balance:		
	Nov-Jan [mm month^{-1}]	-4.24	$+10.69$
	Feb-Apr [mm month^{-1}]	-20.7	$+12.4$
	May-Jun [mm month^{-1}]	-1.84	$+3.3$
	Jul-Oct [mm month^{-1}]	-0.91	$+0.92$
SEBAL evaporation maps	Evaporation per timestep	$\mu \pm 0.3\mu$
	Total dry-season evaporation	$\mu \pm 0.1\mu$
	Parameter: S_{max} [mm]		
	Riverine	500	650
	Dambos	275	500
	Forested	1300	2000
	Highlands	1400	2000
	Parameter: l_p [-]		
	Riverine	0.75	1
	Dambos	1	1
	Forested	0.25	0.4
	Highlands	0.5	0.6

7.5.2 Validation experiment

It must be clear that the calibration of the rainfall runoff model, shown in the previous chapter and the previous sections, is fully *indirect*, meaning that no concomitant time series of modelled and observed discharges could be generated to perform a classical direct calibration. The validation carried out here is therefore fully independent from any performed calibration and hence it is validation in the true sense: not only have the collected records been left out in the parameter inference process, they offer the ability to perform a direct comparison between modelled and observed discharge, given that we have knowledge of the rating curve and that the rainfall inputs are accurate enough. In the previous chapter, this validation material could not be used properly, because all parameters were considered fully lumped over the catchment, while the validation data

have been collected at an interior position in the catchment (upstream area $\sim$ 75 000 $(km)^2$), which requires a certain degree of spatial distribution of parameters and input.

For the validation experiment, optimal parameter sets have been retained by considering models to be behavioural when they obey all the knowledge gathered in the calibration attempts. The spatial distribution of S_{max} and l_p was controlled by imposing a fuzzy rule on the selection of these parameters (shown as red dotted lines in Fig. 7.5). In this validation, only the most likely range of this fuzzy rule was considered. In Table 7.2, all the criteria are summarised.

To find behavioural parameter sets, the self-adapting differential evolution algorithm (Brest $et\ al.$, 2006) was used. This algorithm includes stochastic elements which ensures randomness in the generation of offspring during optimisation. This guarantees that calibrated parameter sets vary in each calibration trial. One hundred calibrated parameter sets were generated in this way and were used to force the hydrological model from the period September 2002 until August 2008. This ensured a generous spin-up time of 5 years for a proper initialisation of soil moisture storage. The result of the one-hundred discharge realisations at Mfuwe together with the discharges, generated by applying Eq. 7.9 on the observed water levels is given in Fig. 7.9.
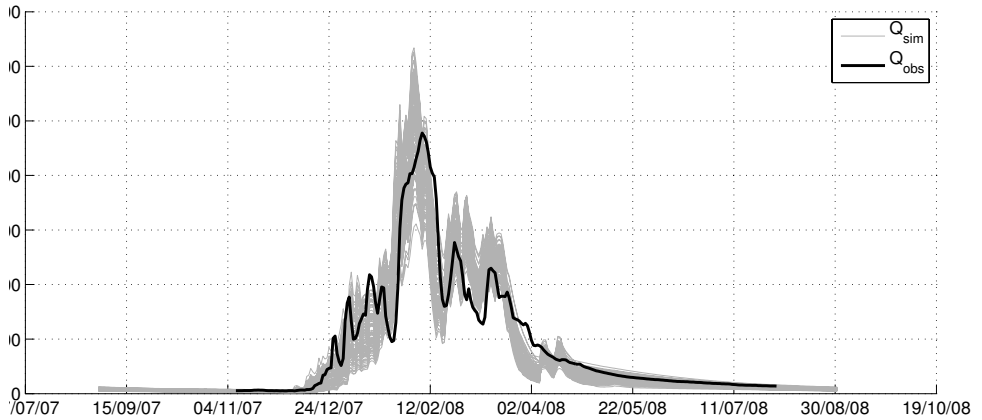
Figure 7.9: Validation of modelled discharge at Mfuwe. The gauged discharges are derived from in-situ measurements of water levels through rating. An estimation of the rating curve has been derived from old records.

Plotting of $h - h_0$ (from the measurements at Mfuwe bridge) and simulated Q, averaged over all 100 simulations on a double log scale (Fig. 7.10), reveals a straight line, which increases my belief in the chosen value for h_0. Hysteretic loops can be observed, the largest
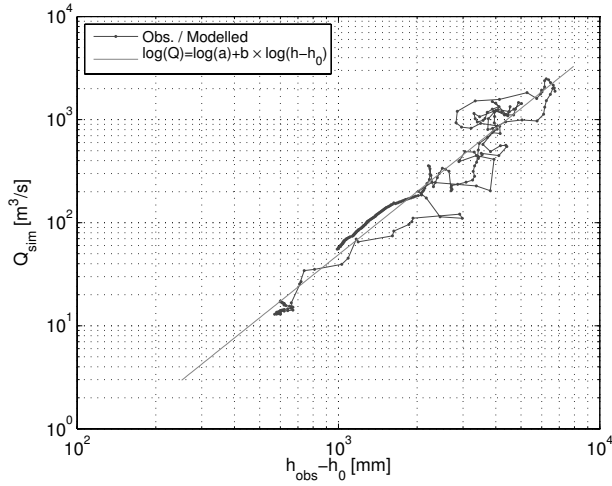
Figure 7.10: Rating relation between simulated discharge (i.e. the average of 100 realisations) and water levels with reference to h_0 at Mfuwe.

being around the flood wave of the beginning of February, which are probably due to the omitting of a short routing time lag. Discrepancies in timing can also be observed in Fig. 7.9, where the observations appear to lag $\sim$1–2 days behind the simulations. In January 2008, a number of small peaks have been missed by the hydrological model. These peaks may have been caused by a tributary, upstream of Mfuwe, which responds faster than other tributaries. The effect of this fast response is likely to attenuate further downstream, which is why this response does not appear in the calibrated models, since their runoff response has been calibrated based on a downstream located data set. Another explanation could be that these first peaks have been generated by very local rainfall events, which have been smoothed out over the larger model units.

It is evident, that the validation is in general successful. It provides independent credibility that the calibration framework, elaborately described in the former two chapters, is valid and useful for predictions in ungauged basins, where data may be scarce, intermittent and non-concomitant, but where we are nonetheless forced to use them. More importantly, the randomly selected behavioural parameter sets produce an envelope of realisations that in most parts of the hydrograph encloses the observed flows. This means that the model is not overconditioned, which is a soft proof that the framework indeed ensures a control on subjectivity.

7.6 Discussion

7.6.1 The value of evaporation information

The assumptions of lumped parameters in the first experiment conducted in this chapter, enabled an unambiguous evaluation of the information content of the evaporation estimates. Although a very simple routine for transpiration has been applied, the results clearly show what type of parameters are expected to be controlled by this information. Parameters related to the soil moisture capacity and the resistance to water stress, previously poorly identifiable, can be well controlled by the evaporation information. It does not mean that the model structure used is the absolute correct model structure for transpiration. In fact, it may be an oversimplification of reality. But this is inherent to the conceptual nature of rainfall runoff models, especially when they are run lumped over such large spatial scales as presented in this case. Given the uncertainty in the information provided, and given the coarse spatial disaggregation applied here, it is not likely that a more complex transpiration routine would give more insight in the information content of the evaporation estimates.

This also underlines that the choice of the model is largely dependent on the application it is meant for. In Chapter 6, Section 6.3 it is stated that this case study intends to 'calibrate a rainfall runoff model, which should be applicable in the estimation of river flows at daily time scales'. This choice has limited the model selection to model structures where the transpiration process is simplified, mainly in order to correctly simulate the soil moisture state variability, which in turn determines to a large extent the non-linearity of the runoff response. One could argue that this oversimplification is indeed problematic when predicting the effect of changing conditions. The oversimplification could then lead to wrongly predicted transpiration, which in turn leads to wrongly predicted moisture states. Therefore, the model may be suitable for the prediction of river flows at daily time scale, but may not be suitable for other purposes. It remains a user choice, whether the functions applied are complex enough for the envisaged application.

7.6.2 Reliability of the model structure

This leads to the next point of discussion. Although the validation gave promising results, there is still a point of concern. How certain are we that the model is indeed robust under changing conditions? This is very difficult to test as it would imply that we should monitor this river basin until these conditions indeed change and should repeatedly confront our model results with the observations while the conditions change. Nonetheless, results from the case study could be used to learn a number of things. For instance, I argued that S_{max} may have unrealistically high values, given the evaporation estimates, especially in

the highland and forest model units. This may in fact point in the direction of deficiencies in the model structure: nature developed the deep-rooting forests to be drought resistant, which is why they are capable of tapping from groundwater sources during the dry season, which may even be replenished laterally. These are processes that have not been included in the model structure and which could lead to better process understanding and therefore an improved robustness of the model under changing conditions. Another example (not concluded from this study) is the fact that *Miombo* leaf shed may be temperature related, rather than moisture related, which would lead to serious discrepancies between model and reality when a temperature shift is imposed by climate change. These types of misconceptions are caused by the poor collaboration between hydrologists and ecologists, where hydrologists are far too keen on the fitting of a hydrograph and not on the improvement of process understanding and synergy between hydrology and ecology.

7.6.3 The curse of equifinality

A last question to address is the remaining equifinality problem. I believe that equifinality still manifests itself in two ways. The first is the selection of spatial disaggregation. This is a highly subjective choice, but may for many applications be crucial. A different disaggregation may reveal slightly different parameter distributions and different final results. The second is the constructed parameter ranges of S_{max} and l_p. Due to the spatial distribution of parameters in space, there is still a wide choice of parameter sets within the fuzzy acceptability rules, which leads to the same answer. The main problem is that S_{max} and l_p in one hydrotope, have exactly the same function in the model structure as S_{max} and l_p in another hydrotope. This leads to compensation between poorly chosen values of S_{max} and l_p in the different hydrotopes, which cannot be straightforwardly resolved. The validation study is therefore not aimed at one specific parameter set but at multiple equally likely parameter sets.

7.7   Conclusions

The calibration framework, presented in Chapter 6 has been proven to be successful in obtaining considerably constrained parameter probability distributions of hydrological models in ungauged basins by means of any available data and with a considerable control of subjectivity. In this chapter, additional soft information was retrieved from a remotely sensed evaporation dataset. This information was shown to force behavioural models to follow the total dry-season evaporation, but also the temporal pattern of transpiration. In an experiment with only lumped parameters, it turned out that this information may

for a great deal remove compensation issues between water balance related parameters, in the Luangwa case study being represented by parameters S_{max}, l_p and B, leaving only a clear interdependency between l_p and B. Accounting for these observations, drastically reduces the variability in possible realisations of the evaporation regime, while additional constraints on the discharge regime are marginal. It is hypothesized here that the reduction in possible realisations of the full dynamics of the model (i.e. not only stream flow), will result in a smaller predictive uncertainty when predicting the effect of future changing conditions such as land cover and climate change.

The spatially distributed nature of evaporation has been used to disaggregate parameter values over dominant land covers. Four different hydrotopes were delineated on the basis of similar evolution of the Normalised Difference Vegetation Index through time. Constraining parameters S_{max} and l_p for each model unit independently, resulted in consistent optimal parameter ranges in the response surface for each dominant land cover, which could be used to construct a fuzzy classification of best performing parameter ranges. The consistent optimal ranges of the response surface are hydrologically meaningful in the sense that they concur with what one can expect from the land surface: landscapes covered by deep rooting vegetation reveal high optimal values for S_{max}, the unsaturated zone capacity, and areas with shallow rooting vegetation, dambos and riverine areas, show much smaller values for S_{max}. Furthermore, the dambo-dominated areas, typically covered with seasonal grasslands, are easily stressed for moisture, which corresponds with high values for l_p.

Validation of a hundred randomly chosen behavioural parameter realisations with a concomitant time series of ground-corrected satellite rainfall estimates and rated discharges showed a remarkable agreement. This validation not only indicates that the calibration is successful but also that we have not overconditioned our model. This implies that the used limits of acceptability have been effective in controlling the subjectivity inherent in parameter uncertainty estimation methods such as GLUE. The presented framework is believed to be useful for many ungauged catchments with uncertain and intermittent data availability.

Chapter 8

Discussion and conclusions

8.1   The value of orthogonal observations

In this thesis, efforts have been made to develop top-down approaches to include some of the most promising new information sources in the development of hydrological models in ungauged basins. Before going into detail, I would like to recall the three main information sources that have been explored in this thesis: ground measurements, satellite data and expert knowledge.

First of all, in-situ measurements have been of great value. In Chapter 5 they allowed for calibration of hydrological models, where only expert knowledge was available to develop a perceptual model and model structure. More importantly, the value and information content of discharge has been shown to be often underrated as hydrologists tend to use one-size-fits-all evaluation measures such as the Nash-Sutcliffe coefficient, which may not be applicable when data is highly uncertain, biased or non-concomitant. Chapter 6 shows that in these conditions, valuable information may be retrieved from any available discharge time series when using innovative measures for model evaluation. These measures should be constructed in such a way that shortcomings of the data at hand do not manifest themselves dominantly and that long time series, if available, are optimally used.

The satellite information used in this study includes gravity observations from the Gravity Recovery and Climate Experiment (GRACE), satellite-based evaporation estimates and satellite-based rainfall estimates. In Chapter 4, the shortcomings of GRACE with respect to hydrological modelling have been analysed and a framework has been proposed to deal with these shortcomings and optimise the information retrieval of GRACE. Chapter 5 describes a successful application of this framework, which fully combines expert knowledge, ground observations and gravity observations. GRACE enables the rejection of hydrological models on the basis of a completely different feature (i.e. the storage variability in time) than traditionally used stream flow records. This justifies the construction of a

more complex hydrological model structure, despite the fact that a basin may be largely ungauged. A similar model feature is evaporation. Especially in semi-arid basins, evaporation accounts for the dominant part of the outgoing fluxes, which means it can provide a strong control on hydrological models. This was demonstrated in Chapter 7, where the possible process realisations were seriously reduced by means of evaporation data. Furthermore, land surface related parameters were spatially distributed into four dominant land cover response classes, based on these data.

Last but not least, expert knowledge provides great insight and perhaps one of the strongest constraints in the choice of a model structure, as it may provide guidelines to what processes may be dominant and how they are interlinked. It was already argued in Chapter 5 that probably many hydrological model structures could have encompassed both discharge data and GRACE data, however the ones that also encompass expert knowledge may be limited to a few.

8.2   A framework to use GRACE for model evaluation: the synergy between hydrologists and geodesists

One of the targets of this thesis is to study the possibility to use GRACE in hydrological modelling of river basins. In Chapter 4, illustrations have been given to emphasize the type of uncertainties and problems that we should expect when dealing with GRACE in hydrology. It appeared that the problem of leakage, i.e. errors in the storage estimates due to the mandatory spatial filtering, may be addressed by a joint effort between hydrologists and geodesists. To this end, a framework has been developed that deals with the following limitations:

- GRACE has a limited spatial resolution. Using GRACE for calibration of hydrological models therefore remains difficult, because the spatial distribution of hydrological processes and climatological inputs is too large to allow for a fully lumped hydrological model at the scale where GRACE observations become meaningful. Conversely, calibration of a hydrological model with spatially distributed parameters on a basin-integrated storage estimate would undoubtedly lead to a great deal of equifinality.
- Spatial filtering affects amplitude and removes not only spatial, but also temporal (even inter-annual) variability. Reduction of information content due to leakage should be considered.

A framework was proposed in Chapter 4 that deals with these limitations. In short, it consists of the following steps: (a) a perceptual and consequently conceptual model is

required, based on prior perceptions of the dominant processes and how they are inter-linked. In order to be able to estimate the bias as a manifestation of leakage, this model should have a satisfactory spatial distribution and should spatially be expanded beyond the region for which it is intended. This conceptual model may be calibrated, using any data available, other than GRACE storage estimates; (b) The first approximation of the hydrological processes within and in the surroundings of the basin may henceforth be employed to make a spatially distributed and time-variable estimate of the storage in the river basin and its surroundings. This estimate should be used to estimate the bias from leakage, which leads to bias-corrected GRACE storage estimates; (c) The hydrological model outputs may consequently be confronted with this bias-corrected storage estimate after which the modeller may decide to improve the model structure and reperform steps a) until c). As long as a fully quantified uncertainty of GRACE data is not available, it is up to the modeller to consider the softness of GRACE as a data source and to decide when this iteration process should stop.

8.3 On the effectiveness of GRACE for model structure inference

In Chapter 5 the proposed framework has been applied in the development of a hydrological model for the upper Zambezi (upstream of Victoria Falls). First, a hydrological model has been constructed, based on prior knowledge collected during a field survey. From the field work, understanding of the dominant processes, storage compartments and their interlinkage was obtained. Through hydrological reasoning, we came to the conclusion that wetlands are dominant factors in creating connectivity between streams and that they function as a fill-and-spill system. The limited availability of old monthly discharge time series helped to test several model structures that were based on the perceptual model, formulated following the observations from the field survey.

The model concept has been applied over the whole of Southern Africa to be able to generate a leakage error estimate. The large spatio-temporal variability in the hydrological cycle resulted in a large amount of error, even with an increasingly large target area. We may conclude that this error significantly reduced the information content retrieved from GRACE: it turned out that leakage not only reduces the average annual amplitude of storage but also reduces interannual variability present in the GRACE signal. According to the GRACE storage estimates, corrected for leakage error, our model simulated both annual and interannual variability remarkably well in the upper Zambezi. Apparently the chosen model structure was able to reproduce the variability caused at interannual time

scales, which allowed for maintaining the chosen model structure as a plausible realisation of the hydrological processes in the upper Zambezi.

From the case study in the upper Zambezi, we may conclude that GRACE can in fact assist in the formulation of hydrological models. By means of this data source, the modeller may gain confidence that the model structure indeed represents the hydrology, or may decide to revise the model structure. The provisos are that the basin's surroundings should be included in the modelling exercise in order to estimate bias due to leakage and that model structural enhancement should be performed iteratively, at least until some degree of convergence between GRACE and modelled storage is reached.

8.4   So how important is 'art' in the modelling process?

An additional important conclusion from this study is the fact that the model was able to reproduce not two, but three important orthogonal data sources: the first two are measured discharges within the catchment and GRACE time variable storage. The third data source may from a mathematical point of view be regarded as vague and ill-quantifiable and is therefore often disregarded in the process of model evaluation. It is the expert knowledge, collected during field surveys which is in my opinion of crucial importance in model structure inference as it is the ultimate validation of a model structure. It can in fact be a justifiable reason for modelling with more parameters and a higher spatial discretisation than equifinality, based on mere numbers, would allow for. An example is availability of local knowledge about vegetation, appearance of wetlands or groundwater level responses to rainfall. Even if these are rough estimates (consider for instance the observed groundwater levels in wetlands throughout the upper Zambezi at the end of the dry season, described in Section 5.3), such information may provide the modeller with crucial constraints on the possible model structures. This could be seen as a manifestation of 'art' in the modelling process (Savenije, 2009). Although perhaps not important for short-term (e.g. flood) predictions, the 'degree of correctness' of a model structure, although by numbers incommensurable, is important when boundary conditions of the model change beyond the calibration conditions, which is the case when land use or climate changes.

8.5   Dealing with subjectivity

The previous conclusions show that satellite information offers a promising perspective towards improving understanding of hydrological processes and henceforth, the conditioning of hydrological model structures and parameters. Nonetheless, remote sensors cannot

observe the river itself. When rainfall-runoff models are concerned, stream flow responses cannot be measured from space, except perhaps on the very large spatial and temporal scale with altimetry instruments. It is therefore in scarcely gauged basins important to be able to profit from any ground data that may be available in a study area. In Chapter 6, a method for model calibration, based on the Generalized Likelihood Uncertainty Estimation (GLUE) has been proposed that suggests to use any available information, be it uncertain, intermittent and non-concomitant with other data sources, needed for environmental models (note that the framework is not limited to rainfall-runoff models alone). In such cases, a modeller has to be prepared to move away from one-size-fits-all likelihood measures such as the Nash-Sutcliffe coefficient (Nash & Sutcliffe, 1970). Instead, the modeller has to define ingenuous informal likelihood measures, which optimally extract information from these data by focussing on signatures, which may manifest a low sensitivity for the poor quality of the data. The appropriateness and value of possible measures used, is in such cases dependent on the modellers perception of data quality and dominant hydrological processes in the studied catchment. In the case study presented in this thesis, the chosen signatures (related to spectral properties, recession curve features and water balance estimates from a benchmark model) were typically relatively insensitive to the direct accuracy of the rainfall inputs. They were deliberately chosen in this way, because it was known a priori that satellite rainfall estimates, used during calibration, were inaccurate in absolute sense, but were expected to be consistent in their spectral properties.

Furthermore, it has been shown that in typical ungauged conditions, such as in the Luangwa basin, many informal likelihood measures may have to be defined to reach useful model conditioning. These may independently have a limited constraining power. Researchers that are unaware of ungauged conditions may not be open minded in accepting that such likelihood measures may be very powerful in ungauged cases (see for instance the discussion on GLUE by Mantovan & Todini, 2006; Beven *et al.*, 2008). They may be seen as a weak, subjective path towards model constraints. In my opinion, there is no such thing as a fully objectively defined likelihood function and therefore this objectivity should not be seen as the fundament of uncertainty estimation. The fact that the school of formal statistics blames the informal school for incoherence (Mantovan & Todini, 2006) is therefore in my opinion first of all a case of the pot calling the kettle black. Even if the conditions are gauged, the formulation of an error model for data, used to drive or calibrate a hydrological model, is as much a subjective choice as the formulation of informal likelihood measures. A posteriori testing of the assumptions underlying the error model may be possible and fruitful, but the modeller has to keep in mind that the error model chosen may as much be subject to equifinality as the hydrological model itself. This could

mean (although the author has not performed any detailed study to support this) that
under changing conditions, a wrongful error model may turn out to give false estimates
of uncertainty. What is more troubling, is the fact that researchers are apparently not
willing to accept that the suitability of uncertainty estimation methods is case specific,
i.e. dependent on the data available and the purpose of the model. Any likelihood measure should therefore be allowed to be combined to find better parameter constraints, as
long as the modeller has used his common sense whether the information provided is not
spurious and has done his best to define objective limits of acceptability.

Here we arrive at a more fundamental problem: how to choose appropriate limits of acceptability? If the modeller decides to move away from formal statistics, then he should
decide when a model structure or parameter set is to be rejected. These are generally
subjective decisions, typically performed in applications of GLUE, that may seriously influence the results. In the case of multiple, relatively low informative likelihood measures,
these choices will result in a multiplication of intermediate parameter distributions that
individually are strongly controlled by subjectivity. This has serious implications on the
validity of the final parameter distribution. The choice of appropriate limits of acceptability, based on a signature rather than a classical likelihood function that summarises
model performance, opens doors to a less subjective way to include information in the
parameter inference, while reducing the risk of overconditioning. This is because these
signatures may have more physical meaning (e.g. the slope of the recession curve) than
a classical likelihood measure. It means that their limits of acceptability may be derived
based on the modellers perception, when they cannot be derived from available data.
When enough data is available, they may be based on for instance the year-to-year variability of the signature, found in these data. What subjectivity remains is limited to the
user's perception of the representativity of the available data for the possible conditions
in the basin (e.g. the user should ask himself if the data encompasses the extremes in the
response) and the selection of a combination of signatures itself.

8.6   The use of multi-observation calibration: evaporation as additional information

Application of the limits of acceptability framework on a discharge time series in the
Luangwa river, led to the inference of a distribution of model parameter sets, which clearly
revealed where in the parameter space constraints were lacking and consequently, what
information could improve parameter constraints. It was clear that parameters, related
to the water balance, were still poorly constrained. Until this point in the analysis, all

signatures used were derived from discharge time series. It is well known that conspiracies between water balance determining parameters exist, when it comes to simulation of discharge. These parameters control the separation of rainfall into streamflow and other processes, and since all signatures used were somehow related to the discharge process, this resulted in poor identifiability of water balance determining parameters. Introduction of evaporation information resulted in an improvement of identifiability, especially of the root zone depth and moisture stress on transpiration, shown in Chapter 7.

It is attractive to use the collected information to relax the spatial distribution of parameters, if the collected data is indeed distributed. This has been shown in Chapter 7, where parameters, strongly related to the evaporation regime in the dry season, were spatially conditioned, based on prior identification of land cover classes. At the basin scale, spatial distribution was still difficult, because wrongfully chosen parameters tend to compensate for each other in the different land cover classes. Nonetheless, a fuzzy classification of the values of spatially distributed parameters could be made. The classification of parameters was hydrologically meaningful in the sense that the differences in the optimal ranges of the parameters between the different land covers could be expected. The fuzzy limits allowed us to find an ensemble of model realisations, that both obey the fuzzy limits and all other limits of acceptability imposed on the outputs, while still considering the uncertainty in the parameters.

The only way to prove that the developed approach, which enables the use of multi-observational calibration within the knowledge that data are uncertain, actually works, is validation with a concomitant data set. The validation in Chapter 7 was performed with the best rainfall estimates at hand, namely a combination of rain gauge measurements and satellite rainfall, and a set of water level measurements collected in the field. The randomly drawn behavioural model realisations most of the time encompassed the stream flow observations. This provided a strong indication that the framework is successful. It combines uncertain, intermittent, non-concomitant, yet extremely valuable old ground records with new remotely sensed data, while considering the uncertainties of all the information derived. Data scarcity will evidently remain a problem in many parts of the world and it is therefore my hope that this framework will be frequently applied in practice, wherever data scarcity exists.

References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986a). An introduction to the European Hydrologic System Système hydrologique européen, SHE, 1: History and philosophy of a physically based, distributed modelling system. *J. Hydrol.*, *87*, 45-59.

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986b). An introduction to the European Hydrologic System Système hydrologique européen, SHE, 2: Structure of a physically based, distributed modelling system. *J. Hydrol.*, *87*, 61-77.

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., et al. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present). *J. Hydrometeor.*, *4*, 1147-1167.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *FAO Irrigation and Drainage Paper No. 56 - Crop Evapotranspiration* (Tech. Rep.). Food and Agriculture Organization.

Allen, R. G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model. *J. Irrig. and Drain. Engrg.*, *133*(4), 380-394.

Andersen, O. B., Seneviratne, S. I., Hinderer, J., & Viterbo, P. (2005). GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave. *Geophys. Res. Lett.*, *32*, L18405.

Arkin, P. A., & Meisner, B. N. (1987). The relationship between large-scale convective rainfall and cold cloud over the Western hemisphere during 1982-84. *Mon. Wea. Rev.*, *115*, 51-74.

Atkinson, S., Woods, R. A., & Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing timescales. *Wat. Resour. Res.*, *38*(12), 1314.

Bagayoko, F. (2006). *Impact of land-use intensity on evaporation and surface runoff: processes and parameters for Eastern Burkina Faso, West Africa*. PhD thesis, Delft, University of Technology.

Bastiaansen, C. J. M. (1995). *Lui River Valley model and some of its applications* (Tech. Rep.). Wageningen University and Research.

Bastiaanssen, W. G. M., Ahmed, M. D., & Chemin, Y. (2002). Satellite surveillance of evaporative depletion across the Indus basin. *Water Resour. Res.*, *38*, 1273-1282.

Bastiaanssen, W. G. M., & Chandrapala, L. (2003). Water balance variability across Sri Lanka for assessing agricultural and environmental water use. *Agric. Wat. Man.*, *58*, 171-192.

Bastiaanssen, W. G. M., Hoekman, H. H., & Roebeling, R. A. (1994). A methodology for assessment of surface resistance and soil water storage variability at mesoscale based on remote sensing measurements. In *IAHS special publication* (p. 65). Wallingford: IAHS Press.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998a). A remote sensing surface energy balance algorithm for land (SEBAL). Part 1. Formulation. *J. Hydrol.*, *212/213*, 198-212.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998b). A remote sensing surface energy balance algorithm for land (SEBAL). Part 2. Validation. *J. Hydrol.*, *212/213*, 213-229.

Bastiaanssen, W. G. M., Noordman, E. J. M., Pelgrum, H., Davids, G., Thoreson, B. P., & Allen, R. G. (2005). SEBAL model with remotely sensed data to improve water-resources management under actual field conditions. *J. Irrig. and Drain Engrg.*, *131*, 85-93.

Beven, K. J. (1989). Changing ideas in hydrology - the case of physically based models. *J. Hydrol.*, *105*, 157-172.

Beven, K. J. (2001). *Rainfall-runoff Modelling: The Primer.* John Wiley & Sons, Ltd., Chichester.

Beven, K. J. (2006). A manifesto for the equifinality thesis. *J. Hydrol.*, *320*, 18-36.

Beven, K. J. (2008). On doing better hydrological science. *Hydrol. Process.*, *22*, 3549-3553.

Beven, K. J., & Binley, A. M. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Proc.*, *6*, 279-298.

Beven, K. J., & Freer, J. E. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.*, *249*, 11-29.

Beven, K. J., Smith, P., & Freer, J. E. (2007). Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. *J. Hydrol.*, *338*, 315-318.

Beven, K. J., Smith, P. J., & Freer, J. E. (2008). So just why would a modeller choose to be incoherent. *J. Hydrol.*, *354*, 15-32.

Boyle, D. (2000). *Multicriteria calibration of hydrological models* (PhD thesis). Dept. of Hydrology and Water Resources, University of Arizona, Tucson.

Brest, J., Greiner, S., Bošković, B., Mernik, M., & Žumer, V. (2006). Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. *IEEE Trans. Evol. Comput.*, *10*(6), 646-657.

Brutsaert, W. (1982). *Evaporation into the atmosphere* (G. T. Csadany, Ed.). Dordrecht: Reidel.

Budyko, M. I. (1974). *Climate and life* (D. H. Miller, Ed.). Academic Press.

Bulygina, N., McIntyre, M., & Wheater, H. (2009). Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis. *Hydrol. Earth Syst. Sci. Discuss.*, *6*, 1907-1938.

Chambers, D. P. (2009). Calculating trends from GRACE in the presence of large changes in continental ice storage and ocean mass. *Geophys. J. Int.*, *176*, 415-419.

Chidumayo, E. N. (2005). Effects of climate on the growth of exotic and indigenous trees in central Zambia. *J. Biogeogr.*, *32*, 111-120.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., et al. (2003). Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.*, *108*(D22), 8842.

Crowley, J. W., Mitrovica, J. X., Bailey, R. C., & Tamisiea, M. E. (2006). Land water storage within the Congo basin inferred from GRACE satellite gravity data. *Geophys. Res. Lett.*, *33*, L19402.

De Groen, M. M., & Savenije, H. H. G. (2006). A monthly interception equation based on the statistical characteristics of daily rainfall. *Wat. Resour. Res.*, *42*, W12417.

De Vries, J. J., Selaolo, E. T., & Beekman, H. E. (2000). Groundwater recharge in the Kalahari, with reference to paleo-hydrologic conditions. *J. Hydrol.*, *238*, 110-123.

Di Baldassarre, G., & Montanari, A. (2008). Uncertainty in river discharge observations: a quantitative analysis. *Hydrol. Earth Syst. Sci. Discuss.*, *6*, 39-61.

Didszun, J., & Uhlenbrook, S. (2008). Scaling of dominant runoff generation processes: Nested catchments approach using multiple tracers. *Wat. Resour. Res.*, *44*, W02410.

Farah, H. O. (2001). *Estimation of regional evaporation under different weather conditions from satellite and meteorological data; A case study in the Naivasha Basin, Kenya.* PhD thesis, Wageningen University.

Fenicia, F., McDonnell, J. J., & Savenije, H. H. G. (2008a). Learning from model improvement: On the contribution of complementary data to process understanding. *Wat. Resour. Res.*, *44*, W06419.

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2006). Is the groundwater reservoir linear? Learning from data in hydrological modelling. *Hydrol. Earth Syst.*

Sci., *10*, 139-150.

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2008b). Understanding catchment behavior through stepwise model concept improvement. *Wat. Resour. Res.*, *44*, W01402.

Ferraro, R. R., & Marks, G. F. (1995). The development of SSM/I rain-rate retrieval algorithms using ground-based radar measurements. *J. Atmos. Oceanic Technol.*, *12*, 755-770.

Franks, S. W. (2007). Integrating models, methods and measurements for prediction in ungauged basins. In D. Schertzer, P. Hubert, S. Koide, & K. Takeuchi (Eds.), *Proceedings of the PUB kick-off meeting, brasilia, november 2002* (Vol. 309, p. 13-21). England.

Franks, S. W., Gineste, P., Beven, K. J., & Merot, P. (1998). On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resour. Res.*, *34*, 787-797.

Freer, J. E., McMillan, H., McDonnell, J. J., & Beven, K. J. (2004). Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures. *J. Hydrol.*, *291*, 254-277.

Frost, P. (1996). The Miombo in Transition: Woodlands and Welfare in Africa. In B. M. Campbell (Ed.), (p. 11-58). Center for International Forestry Research.

Fuller, D. O. (1999). Canopy phenology of some Mopane and Miombo woodlands in eastern Zambia. *Global Ecol. Biogeogr.*, *8*(3-4), 199-209.

Gammelsrod, T. (1996). Effect of Zambezi River management on the prawn fishery of the Sofala Bank. In M. C. Acreman & G. E. Hollis (Eds.), *Water management and wetlands in Sub-Saharan Africa* (p. 199-224). IUCN.

Gragne, A. S., Uhlenbrook, S., Mohamed, Y., & Kebede, S. (2008). Catchment modeling and model transferability in upper Blue Nile basin, lake Tana, Ethiopia. *Hydrol. Earth Syst. Sci. Discuss.*, *5*, 811-842.

Grayson, R., & Blöschl, G. (2001). *Spatial patterns in catchment hydrology: observations and modelling.* Cambridge University Press, Cambridge.

Güntner, A. (2008). Improvement of global hydrological models using GRACE data. *Surv. Geophys.*, *29*(4-5), 375-397.

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Process.*, *22*, 3802-3813.

Han, S. C., Jekeli, C., & Shum, C. K. (2004). Time-variable aliasing effects of ocean tides, atmosphere, and continental water mass on monthly mean GRACE gravity field. *J. Geophys. Res.*, *109*, B04403.

Hasan, S. (2009). *Terrestrial water storage change from temporal gravity variation.* PhD thesis, Wageningen University and Research.

Herbst, M., Gupta, H. V., & Casper, M. C. (2009). Mapping model bahaviour using Self-Organizing Maps. *Hydrol. Earth Syst. Sci., 13*, 395-409.

Herman, A., Kumar, V. B., Arkin, P. A., & Kousky, J. V. (1997). Objectively determined 10-day African rainfall estimates created for Famine Early Warning Systems. *Int. J. Remote Sensing, 18*(10), 2147-2159.

Hoekstra, A. Y., Savenije, H. H. G., & Chapagain, A. K. (2001). An integrated approach towards assessing the value of water: a case study on the Zambezi basin. *Integrated Assessment, 2*(4), 199-208.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor., 8*, 38-55.

Hughes, D. A., Andersson, L., Wilk, J., & Savenije, H. H. G. (2006). Regional calibration of the Pitman model for the Okavango river. *J. Hydrol., 30*(1-2), 30-42.

IGBP. (1990). *The initial core projects* (Report No. 12). IGBP.

Immerzeel, W. W., & Droogers, P. (2008). Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol., 349*, 411-424.

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Phil. Trans. R. Soc. Lond., 273*, 593-610.

Jia, L., Su, Z., Van den Hurk, B. J. J. M., Menenti, M., Moene, A. F., Bruin, H. A. R. D., et al. (2003). Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ATSR measurements. *Phys. Chem. Earth, 28*, 75-88.

Johrar, R. K. (2002). *Estimation of effective soil hydraulic parameters for water management studies in semi-arid zones.* PhD thesis, Wageningen University.

Kelly, K. S., & Krzysztofowicz, R. (1997). A bivariate meta-Gaussian density for use in hydrology. *Stoch. Hydrol. Hydraul., 11*(1), 17-31.

Klees, R., Revtova, E. A., Gunter, B. C., Ditmar, P., Oudman, E., Winsemius, H. C., et al. (2008). The design of an optimal filter for monthly GRACE gravity models. *Geophys. J. Int., 175*(2), 417-432.

Klees, R., Zapreeva, E. A., Winsemius, H. C., & Savenije, H. H. G. (2007). The bias in GRACE estimates of continental water storage variations. *Hydrol. Earth Syst. Sci., 11*, 1227-1241.

Klokočnik, J., Wagner, C. A., Kostelecký, J., Bezděk, A., Novák, P., & McAdoo, D. (2008). Variations in the accuracy of gravity recovery due to ground track variability:

GRACE, CHAMP, and GOCE. *J. Geod.*, *82*, 917-927.

Kongo, V. M., Kosgei, J. R., Jewitt, G. P. W., & Lorentz, S. A. (2007). Establishment of a catchment monitoring network through a participatory approach in a small rural catchment in South Africa. *Hydr. Earth Syst. Sci. Discuss.*, *4*, 3793-3837.

Krzysztofowicz, R., & Kelly, K. S. (2000). Hydrologic uncertainty processor for probabilistic river stage forecasting. *Wat. Resour. Res.*, *36*(11), 3265-3277.

Kuczera, G., Kavetski, D., Franks, S., & Thyer, M. (2006, November). Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *J. Hydrol.*, *331*(1-2), 161-177.

Lamb, R., & Beven, K. J. (1997). Using interactive recession curve analysis to specify a general catchment storage model. *Hydrol. Earth Syst. Sci.*, *1*, 101-113.

Lee, H., Sivapalan, M., & Zehe, E. (2005). Representative elementary watershed (REW) approach, a new blueprint for distributed hydrologic modelling at the catchment scale: development of closure relations. In C. Spence, J. W. Pomeroy, & A. Pietroniro (Eds.), (p. 165-218). Canadian Water Resources Association (CWRA).

Lee, H., Zehe, E., & Sivapalan, M. (2007). Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the CREW model. *Hydrol. Earth Syst. Sci.*, *11*(2), 819-849.

Lewis, D. M. (1991). Observations of tree growth, woodland structure and elephant damage on Colophospermum mopane in Luangwa valley, Zambia. *Afr. J. Ecol.*, *29*, 207-221.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergstrom, S. (1997). Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.*, *201*, 272-288.

Liu, X. (2008). *Global gravity field recovery from satellite-to-satellite track data with the acceleration approach.* PhD thesis, Delft, University of Technology, Optima, Graphical Communication, the Netherlands.

Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward and integrated data assimilation framework. *Wat. Resour. Res.*, *43*, W07401.

LSA SAF. (2007). *Validation report* (Tech. Rep. No. SAF/LAND/IM/VR/1.6). Lisboa: LSA SAF.

Makinwa, K. A. A., & Huijsing, J. H. (2002). A smart wind sensor using thermal sigma-delta modulation techniques. *Sens. Actuators, A*, *97-98*, 15-20.

Makkink, G. F. (1957). Testing the Penman formula by means of lysimeters. *Int. J. Water Eng.*, *11*, 277-288.

Mantovan, P., & Todini, E. (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *J. Hydrol.*, *330*, 368-381.

McGlynn, B. L., McDonnell, J. J., & Brammer, D. B. (2002). A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. *J. Hydrol.*, *257*, 1-26.

Merz, R., & Blöschl, G. (2008a). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resour. Res.*, *44*, W08432.

Merz, R., & Blöschl, G. (2008b). Flood frequency hydrology: 2. Combining data evidence. *Water Resour. Res.*, *44*, W08433.

Messer, H., Zinevich, A., & Alpert, P. (2006). Environmental monitoring by wireless communication networks. *Science*, *312*, 713.

Mohamed, Y. A., Bastiaanssen, W. G. M., & Savenije, H. H. G. (2004). Spatial variability of evaporation and moisture storage in the swamps of the upper Nile studied by remote sensing techniques. *J. Hydrol.*, *289*, 145-164.

Montanari, A. (2007). What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrol. Process.*, *21*, 841-845.

Montanari, A., & Brath, A. (2004). A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Wat. Resour. Res.*, *40*, W01106.

Montanari, A., & Toth, E. (2007). Calibration of hydrological models in the spectral domain: an opportunity for scarcely gauged basins? *Wat. Resour. Res.*, *43*, W05434.

Monteith, J. L. (1981). Evaporation and surface temperature. *Q. J. R. Meteorol. Soc.*, *107*(451), 1-27.

Mul, M. L. (2009). *Understanding hydrological processes in an ungauged catchment in sub-Saharan Africa*. PhD thesis, UNESCO-IHE, Taylor & Francis Group plc, London, UK.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models, part 1: A discussion of principles. *J. Hydrol.*, *10*, 282-290.

New, M., Lister, D., Hulme, M., & Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Res.*, *21*, 1-25.

Ngo-Duc, T., Laval, K., Ramillien, G., Polcher, J., & Cazenave, A. (2007). Validation of the land water storage simulated by Organising Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) with Gravity Recovery and Climate Experiment (GRACE) data. *Wat. Resour. Res.*, *43*, W04427.

Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., Gulden, L. E., & Su, H. (2007). Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. *J. Geophys. Res.*, *112*, D07103.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., & Le Moine, N. (2008). Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *J. Hydrol.*, *44*, W03413.

Partijs, M. A. P., Makinwa, K. A. A., & Huijsing, J. H. (2005). A CMOS smart temperature sensor with a 3σ inaccuracy of $\pm 0.1°C$ from -55°C to 125°C. *IEEE J. Solid-state Circuits*, *40*, 2805-2815.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London*, *A193*, 120-145.

Pitman, W. V. (1973). *A mathematical model for generating monthly river flow from meteorological data in Southern Africa* (Report No. 2). University of Witwatersrand, Dept. of Civil Eng., Hydrol. Res. Unit.

Pokhrel, P., Yilmaz, K. K., & Gupta, H. V. (in press). Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. Hydrol.*

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Wea. Rev.*, *100*, 81-92.

Ramillien, G., Frappart, F., Güntner, A., Ngo-Duc, T., Cazenave, A., & Laval, K. (2006). Time variations of the regional evapotranspiration rate from Gravity Recovery And Climate Experiment GRACE satellite gravimetry. *Wat. Resour. Res.*, *42*, W10403.

Reggiani, P., Sivapalan, M., & Hassanizadeh, S. M. (1998). A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Adv. Water Resour.*, *22*, 367-398.

Reynolds, R. W. (1988). A real-time global sea surface temperature analysis. *J. Climate*, *1*, 75-86.

Rijtema, P. E., & Aboukhaled, A. (1975). *Crop water use. in: Aboukhaled et al. (eds). research on crop water use, salt affected soils and drainage in the Arab Republic of Egypt* (Tech. Rep.). FAO, Near East Regional Office.

Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J., & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi river basin (USA) using GRACE. *Hydrogeol. J.*, *15*(1), 159-166.

Rodell, M., & Famiglietti, J. S. (2002). The potential for satellite-based monitoring of groundwater storage changes using GRACE: the High Plains aquifer, Central US. *J. Hydrol.*, *263*, 245-256.

Rodell, M., Famiglietti, J. S., Chen, J., Seneviratne, S. I., Viterbo, P., Holl, S., et al. (2004). Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophys. Res. Lett.*, *31*, L20504.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.*, *85*, 381-394.

Rougier, J. (2007). Probabilistic inference for future climate using an ensemble of climate

model evaluations. *Clim. Change*, *81*(3-4), 247-264.

Savenije, H. H. G. (1997). Determination of evaporation from a catchment water balance at a monthly time scale. *Hydrol. Earth Syst. Sci.*, *1*, 93-100.

Savenije, H. H. G. (2001). Equifinality, a blessing in disguise? *Hydrol. Process.*, *15*, 2835-2838.

Savenije, H. H. G. (2004). The importance of interception and why we should delete the term evapotranspiration from our vocabulary. *Hydrol. Process.*, *18*, 1507-1511.

Savenije, H. H. G. (2009). The art of hydrology. *Hydr. Earth Syst. Sci.*, *13*, 157-161.

Schaefli, B., & Gupta, H. V. (2007). Do Nash values have value? *Hydrol. Process.*, *21*, 2075-2080.

Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K.-H., Dahle, C., König, R., et al. (2008). Hydrological signals observed by the GRACE satellite mission. *Surv. Geophys.*, *29*(4-5), 319-334.

Schüttemeyer, D. (2005). *The surface energy balance over drying semi-arid terrain in West Africa*. PhD thesis, Wageningen Universiteit.

Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen, W. G. M., & Bierkens, M. F. P. (2003). Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Adv. Wat. Resour.*, *26*, 151-159.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., & Hutley, L. B. (2008). An optimality-based model of the coupled soil moisture and root dynamics. *Hydrol. Earth Syst. Sci.*, *12*(3), 913-932.

Seibert, J. (2001). On the need for benchmarks in hydrological modelling. *Hydrol. Process.*, *15*, 1063-1064.

Seibert, J., & McDonnell, J. J. (2002). On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Wat. Resour. Res.*, *38*(11), 1241.

Selker, J., Van de Giesen, N. C., Westhoff, M., Luxemburg, W., & Parlange, M. B. (2006). Fiber optics opens window on stream dynamics. *Geophys. Res. Lett.*, *33*, L24401.

Seneviratne, S. I., Viterbo, P., Luthi, D., & Schar, C. (2004). Inferring changes in terrestrial water storage using ERA-40 Reanalysis data: the Mississippi river basin. *J. Climate*, *17*, 2039-2057.

Seo, K.-W., Wilson, C. R., Chen, J., & Waliser, D. E. (2008). GRACE's spatial aliasing error. *Geophys. J. Int.*, *172*, 41-48.

Sivapalan, M. (2003). Prediction in Ungauged Basins: a grand challenge for theoretical hydrology. *Hydrol. Process.*, *17*, 3163-3170.

Sklash, M. G., & Farvolden, R. N. (1979). The role of groundwater in storm runoff. *J. Hydrol.*, *43*, 45-65.

Son, K., & Sivapalan, M. (2007). Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. *Wat. Resour. Res.*, *43*, W01415.

Swenson, S. C., & Wahr, J. (2002). Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. *J. Geophys. Res.*, *107*(B9), 2193.

Swenson, S. C., & Wahr, J. (2006). Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.*, *33*, L08402.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE measurements of mass variability in the earth system. *Science*, *305*, 503-505.

Tapley, B. D., & Reigber, C. (2001). The GRACE mission: status and future plans. In *Eos. trans, fall meeting* (Vol. 82, p. G14C-02).

Uhlenbrook, S., Frey, M., Leibundgut, C., & Maloszewski, P. (2002). Hydrograph separations in a mesoscale mountainous basin at event and seasonal timescales. *Wat. Resour. Res.*, *38*(6), 1-14.

Uhlenbrook, S., & Leibundgut, C. (2002). Process-oriented catchment modelling and multiple-response validation. *Hydrol. Process.*, *16*, 423-440.

Uhlenbrook, S., & Wenninger, J. (2006). Identification of flow pathways along hillslopes using electrical resistivity tomography (ERT). In M. Sivapalan *et al*. (Eds.), *Predictions in ungauged basins: promise and progress* (Vol. 303, p. 15-20). England.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, *131*, 2961-3012.

Vaché, K. B., & McDonnell, J. J. (2006). A process-based rejectionist framework for evaluating catchment runoff model structure. *Wat. Resour. Res.*, *42*, W02409.

Van de Giesen, N., Kunstmann, H., Jung, G., Liebe, J., Andreini, M., & Vlek, P. L. G. (2002). The GLOWA Volta project: Integrated assessment of feedback mechanisms between climate, landuse, and hydrology. In M. Beniston (Ed.), *Climatic Change. Implications for the Hydrological Cycle and for Water Management*.

Vaz, A. C. (2000). Some lessons from the floods in mozambique. In *International conference for Mozambique floods*.

Velicogna, I., & Wahr, J. (2006). Measurements of time-variable gravity show mass loss in Antarctica. *Science*, *311*, 1754-1756.

Vogel, R. M., & Sankarasubramanian, A. (2003). Validation of a watershed model without calibration. *Wat. Resour. Res.*, *39*(10), 1292.

Von der Heyden, C. J., & New, M. G. (2004). Groundwater pollution on the Zambian

Copperbelt: deciphering the source and the risk. *Sci. Total Environ.*, *327*, 17-30.

Voogt, M. (2006). *METEOLOOK, a physically based regional distribution model for measured meteorological variables.* MSc thesis, Delft, University of Technology.

Vose, R. S., Schmoyer, R. L., Steurer, P. M., Peterson, T. C., Heim, R., Karl, T. R., et al. (1992). *The Global Historical Climatology Network: long-term monthly temperature, precipitation, sea level pressure, and station pressure data* (ORNL/CDIAC-53, NDP-041). Oak Ridge, TN: Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., & Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Wat. Resour. Res.*, *39*(8), 1214.

Wagener, T. (2003). Evaluation of catchment models. *Hydrol. Process.*, *17*, 3375-3378.

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stoch. Environ. Res. Risk Assess.*, *19*, 378-387.

Wagener, T., Sivapalan, M., McDonnell, J. J., Hooper, R., Lakshmi, V., Liang, X., et al. (2004). Predictions in Ungauged Basins as a catalist for multidisciplinary hydrology. *EOS. Trans. AGU*, *85*(44).

Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *J. Geophys. Res.*, *103*(B12), 30205-30230.

Wahr, J., Swenson, S. C., & Velicogna, I. (2006). Accuracy of GRACE mass estimates. *Geophys. Res. Lett.*, *33*, L06401.

Wahr, J., Swenson, S. C., Zlotnicki, V., & Velicogna, I. (2004). Time-variable gravity from GRACE: first results. *Geophys. Res. Lett.*, *31*, L11501.

Werth, S., Güntner, A., Petrovic, S., & Schmidt, R. (2009). Integration of GRACE mass variations into a global hydrological model. *Earth and Planetary Science Letters*, *277*(1-2), 166-173.

Westhoff, M. C., Savenije, H. H. G., Luxemburg, W. M. J., Stelling, G. S., Van de Giesen, N. C., Selker, J. S., et al. (2007). A distributed stream temperature model using high resolution temperature observations. *Hydrol. Earth Syst. Sci.*, *11*(4), 1469-1480.

Whittle, P. (1953). Estimation and information in stationary time series. *Ark. Mat.*, *2*(23), 423-434.

Winsemius, H. C., Savenije, H. H. G., & Bastiaanssen, W. G. M. (2008). Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? *Hydrol. Earth Syst. Sci.*, *12*(4), 1403-1413.

Winsemius, H. C., Savenije, H. H. G., Gerrits, A. M. J., Zapreeva, E. A., & Klees, R. (2006). Comparison of two model approaches in the Zambezi river basin with regard

to model confidence and identifiability. *Hydrol. Earth Syst. Sci.*, *10*, 339-352.

Winsemius, H. C., Savenije, H. H. G., Van de Giesen, N. C., Van den Hurk, B. J. J. M., Zapreeva, E. A., & Klees, R. (2006). Assessment of Gravity Recovery and Climate Experiment (GRACE) temporal signature over the upper Zambezi. *Wat. Resour. Res.*, *42*, W12201.

Winsemius, H. C., Schaefli, B., Montanari, A., & Savenije, H. H. G. (under review). On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information. *Wat. Resour. Res.*

Wolski, P., Savenije, H. H. G., Murray-Hudson, M., & Gumbricht, T. (2006). Modelling of the flooding in the okavango delta, botswana, using a hybrid reservoir-GIS model. *J. Hydrol.*, *331*(1-2), 58-72.

Xie, P., & Arkin, P. A. (1995). An intercomparison of gauge observations and satellite estimates of monthly precipitation. *J. Appl. Meteor.*, *34*, 1143-1160.

Xie, P., & Arkin, P. A. (1996). Analyses of global monthly precipitation using gauge observations, satellite estimates and numerical model predictions. *J. Clim.*, *9*, 840-858.

Yadav, M., Wagener, T., & Gupta, H. (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Adv. Wat. Resour.*, *30*, 1756-1774.

Yilmaz, K. K., Gupta, H. V., & Wagener, T. (2008). A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Wat. Resour. Res.*, *44*, W09417.

Zhang, G. P., & Savenije, H. H. G. (2005). Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach. *Hydrol. Earth Syst. Sci. Discuss.*, *2*, 639-690.

Zhang, G. P., Savenije, H. H. G., Fenicia, F., & Pfister, L. (2006). Modelling subsurface storm flow with the Representative Elementary Watershed (REW) approach: application to the Alzette River Basin. *Hydrol. Earth Syst. Sci.*, *10*, 937-955.

Zhao, L., & Weng, F. (2002). Retrieval of ice cloud parameters using the Advanced Microwave Sounding Unit. *J. Appl. Meteor.*, *41*, 384-395.

Acknowledgements

It is really amazing how many people have helped me struggling my way through the unknowns of ungauged basins, to perform my PhD research in general, but also to make sure that there are things that broaden the mind in other ways. My conversations, collaborations and great times with all these people have helped shape my ideas into this thesis and propositions. Below I want to show my gratitude to some of the most important ones.

First of all, the understanding of hydrological processes in the region, especially the thoughts shared with the reader in Chapter 5, could not have been gained without field visits in Zambia. Michael Mutale facilitated our trips, told many stories about the hydrology and water resources of the Zambezi, has drawn my attention towards the Luangwa catchment and finally, shared many evenings with a good glass of wine or beer and intriguing conversations about life, nature and God with me and Huub. Bwana, I cannot thank you enough for all the support and knowledge that you shared with me. Your hospitality (I really feel at home in my 'Lusaka apartment'), open-mindedness and kindness convinces me that we have become close friends to each other.

There have been many others in Zambia that helped me a great deal during my visits, a few of whom I would like to mention in particular. From the department of Water Affairs, I would like to thank Jack Nkhoma and Chris Chileshe. Jack, it is always a pleasure to meet you, and I hope we can go on a field trip once again to the hot Luangwa valley. Chris, many thanks for sharing important data sets with me, allowing me to develop my case studies. I appreciate your hospitality and the opportunities to share my work with Water Affairs associates. I would like to thank Maurice Muchinda and Boniface Mbewe of the Meteorological Department for sharing valuable meteorological data. A special thanks goes to James Milanzi of the Zambia Wildlife Authority. James, your assistance in the rehabilitation of the Mfuwe station has resulted in an extremely valuable validation data set of water levels, which I gratefully used to convince the scientific public. This validation also required rainfall numbers in the field, to yield a reliable input data set. These were provided by a number of people that, besides running a lodge in a remote area, succeed in maintaining valuable rainfall stations. I'd like to thank Mike and Lari Merrett (Mutinondo Wilderness), John Coppinger (Tafika, South Luangwa), Ed Sayer and Claire Lewis (FZS,

Marula Puku, North Luangwa) and Mark Harvey (Kapishya hot springs, Shiwa Ng'andu and Buffalo camp, North Luangwa) for this extremely valuable data. I hope that this thesis shows how important it is to keep collecting rainfall observations in the field.

Back in Europe, I have had the pleasure to collaborate with many knowledgeable people. Roland Klees, Elena Revtova and Brian Gunter helped me shape my ideas of the value of satellite gravity measurements in hydrology. Wim Bastiaanssen, Bart van den Hurk and associates of Waterwatch have considerably improved my knowledge of the energy balance over land and the use of meteorological and satellite data therein. Together with my direct colleague, Bettina Schaefli, we came up with ideas to combine these data in hydrological models. Bettina, I enjoy working together with you. It was you who came up with the idea to use spectral domain information in my ungauged case. After this, we started collaborating with Alberto Montanari, whom I also want to mention here. Alberto, many thanks for all the inspiring skype and e-mail conversations and your hospitality at your department in the Università di Bologna. I am now able to find my way in the fuzzy world of uncertainty and I learned how to write a paper with help of an inspiring professor in more or less five days.

Of course I greatly acknowledge my promotor, Huub Savenije. Huub, it is really amazing how you always manage to make time for me, how inspiring you can be in the field and during our discussions, and how good a friend you can be outside the working environment. I sincerely hope that we may keep on collaborating in the future.

I have had a great time in the department of Water Resources. The informal atmosphere, large amount of social activities, and great colleagues and students really made that I came to work with a smile on my face. Wim, many thanks for all the good times in Luxembourg, during drinks of coffee and beers. Miriam, many thanks for assisting in producing latex code, organising Fuut drinks and all the fun in Luxembourg. Fabrizio, you were a true inspiration in the last years. Tom, thanks for being my office mate, your dry sense of humour kept me smiling, even in busy times. It was also sometimes a relief to have an office mate with a completely different subject to talk about from time to time.

I want to thank my friends who have made sure that not my whole life consisted of doing a PhD. In particular I have been happy to join Choir Delft in the last years where I have met many old and new friends, also in the board, and which allowed me to express me in music and not only in the form of scientific publications. Sjoerd and Laurent, thanks for being close friends during many years. It is really an honour that you wanted to be my paranimfs during one of the most important days of my life. I really look forward to more pleasant times with you in the future. Of course I want to thank all my friends but I would prefer to rather do this in person than in the form of an exhaustive list.

Usually one takes the support of family for granted, but here I would like to mention that their support has been one of the most important things in my development. Without the stimulation of my parents and sisters, I would never have come this far. I also want to thank Jan and Maria. You have given me a lot of support in the past years. Finally of course I thank Hanneke. Naturally, obtaining a PhD degree sounds nice, but in the end, you are the most important part of my life.

Hessel Winsemius

Delft, July 2009

About the author

Hessel Winsemius was born in Amsterdam, the Netherlands on the 21st of January 1980. He studied Civil Engineering at Delft, University of Technology (DUT), where he graduated in 2004 at the chair of hydrology and ecology on a hydrological model for efficient ensemble flood forecasts in the Rhine on the basis of ensemble precipitation forecasts. During his studies, he was active in several committees and in the board of the student music society of Delft, Krashna Musika.

From 2004 until 2009, he worked at the chair of hydrology, department of Water Resources of DUT on a PhD research. He developed methods to use several types of satellite observations as complementary information in the construction, calibration and validation of hydrological models. The methods were specifically developed for scarcely gauged and semi-arid regions, with case studies on several tributaries of the Zambezi river in Southern Africa. During his PhD, he participated in several international conferences and meetings, including the annual EGU assembly, the annual Waternet meeting in Southern Africa and several IAHS meetings. In October 2007, he gave an invited speech during the annual Boussinesq Lecture.

Hessel Winsemius has been active in several lectures of the chair of hydrology and the course 'Water Resources Assessment in Sub-Saharan Africa: Predictions in ungauged and data scarce basins', given each year by the Waternet network in the SADC region. He supervised several MSc students during his PhD work.

He was author and co-author of several papers in peer-reviewed journals, not all of them related to the content of this thesis.

Publications:

Winsemius, H. C., Schaefli, B., Montanari, A., Savenije, H. H. G. (under review). On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information. *Wat. Resour. Res.*

Winsemius, H. C., Savenije, H. H. G., Bastiaanssen, W. G. M. (2008). Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins? *Hydrol. Earth Syst. Sci.*, *12*(4), 1403-1413.

Klees, R., Revtova, E. A., Gunter, B. C., Ditmar, P., Oudman, E., Winsemius, H. C., et al. (2008). The design of an optimal filter for monthly GRACE gravity models. *Geophys. J. Int.*, *175*(2), 417-432.

Fenicia, F., Savenije, H. H. G., Winsemius, H. C. (2008). Moving from model calibration towards process understanding. *Phys. Chem. Earth*, *33*, 1057-1060.

Klees, R., Liu, X., Wittwer, T., Gunter, B. C., Revtova, E. A., Tenzer, R., et al. (2008). A comparison of global and regional GRACE models for land hydrology. *Surv. Geophys.*, *29*(4-5), 335-359.

Hut, R., Ertsen, M., Joeman, N., Vergeer, N., Winsemius, H. C., Van de Giesen, N. C. (2008). Effect of sand storage dams on groundwater levels with examples from Kenya. *Phys. Chem. Earth*, *33*, 56-66.

Klees, R., Zapreeva, E. A., Winsemius, H. C., Savenije, H. H. G. (2007). The bias in GRACE estimates of continental water storage variations. *Hydrol. Earth Syst. Sci.*, *11*, 1227-1241.

Winsemius, H. C., Savenije, H. H. G., Van de Giesen, N. C., Van den Hurk, B. J. J. M., Zapreeva, E. A., Klees, R. (2006). Assessment of Gravity Recovery and Climate Experiment (GRACE) temporal signature over the upper Zambezi. *Wat. Resour. Res.*, *42*, W12201.

Winsemius, H. C., Savenije, H. H. G., Gerrits, A. M. J., Zapreeva, E. A., Klees, R. (2006). Comparison of two model approaches in the Zambezi river basin with regard to model confidence and identifiability. *Hydrol. Earth Syst. Sci.*, *10*, 339-352.