Hydrol. Earth Syst. Sci. Discuss., 5, 2293–2318, 2008
www.hydrol-earth-syst-sci-discuss.net/5/2293/2008/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins?

H. C. Winsemius[1], H. H. G. Savenije[1], and W. G. M. Bastiaanssen[1,2]

[1]Department of Water Management, Delft, University of Technology, Delft, The Netherlands
[2]WaterWatch, Wageningen, The Netherlands

Received: 4 July 2008 – Accepted: 8 July 2008 – Published: 18 August 2008

Correspondence to: H. C. Winsemius (h.c.winsemius@tudelft.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In this study, land surface related parameter distributions of a conceptual semi-distributed hydrological model are estimated by employing time series of satellite-based evaporation estimates during the dry season as explanatory information. A key application for this approach is to identify part of the parameter distribution space in ungauged river basins without the need for ground data. The information, contained in the evaporation estimates implicitly imposes compliance of the model with the largest water balance term, evaporation, and a spatially and temporally realistic depletion of soil moisture within the dry season. Furthermore, the model results can provide a better understanding of the information density of remotely sensed evaporation.

The approach has been applied to the ungauged Luangwa river basin (150 000 $(\mathrm{km})^2$) in Zambia. Model units were delineated on the basis of similar land cover. For each model unit, model parameters for which evaporation is sensitive, have been conditioned on the evaporation estimates by means of Monte-Carlo sampling. The results show that behavioural parameter sets for model units with similar land cover, are indeed clustered. The clustering reveals hydrologically meaningful signatures in the parameter response surface: wetland-dominated areas (also called dambos) show optimal parameter ranges that reflect a relatively small unsaturated zone (due to the shallow rooting depth of the vegetation) and moisture stressed vegetation. The forested areas and evergreen highlands show parameter ranges that indicate a much deeper root zone and drought resistance.

Unrealistic parameter ranges, found for instance in the high optimal field capacity values in the highlands may indicate model structural deficiencies. We believe that in these areas, groundwater uptake into the root zone and lateral movement of groundwater should be included in the model structure. Furthermore, a less distinct parameter clustering was found for forested model units. We hypothesize that this is due to the presence of 2 dominant forest types that differ substantially in their moisture regime. Therefore, this could indicate that the spatial discretization used in this study is over-

simplified.

This constraining step with remotely sensed data is useful for Bayesian updating in ungauged catchments. To this end trapezoidal shaped fuzzy membership functions were constructed that can be used to constrain parameter realizations in a second calibration step if more data becomes available. Especially in semi-arid areas such as the Luangwa basin, traditional rainfall-runoff calibration should be preceded by this step because evaporation represents a much larger term in the water balance than discharge and because it imposes spatial variability in the water balance. It justifies that land surface related parameters are distributed. Furthermore, the analysis reveals where hydrological processes may be ill-defined in the model structure and how accurate our spatial discretization is.

1 Introduction

Hydrological models in data sparse areas are often over-simplified. This is often due to the lack of observational data to justify more complexity. As a result, *parsimony* in model parameters is often advocated (Beven and Binley, 1992; Beven and Freer, 2001; Savenije, 2001) to prevent the undesirable occurrence of equifinality (e.g. Beven and Binley, 1992; Beven and Freer, 2001). Although parsimony results in simple and to a certain extent identifiable models, their predictive capacity, for instance of land cover changes, is rather small, because parameters usually have little physical meaning and cannot represent the variability inherent in the landscape. Even with simple models, parameters are often poorly identifiable (e.g. Uhlenbrook et al., 1999) and cannot be justifiably distributed in space in view of the problem of equifinality. A related issue is model structural uncertainty, which is probably even more difficult to define and quantify (Wagener and Gupta, 2005; Young, 2001). Furthermore, in many remote river basins, especially in developing countries, measurement networks are collapsing. Often only old records (often from colonial periods) exist which cannot be confronted with new (remotely sensed) data sources. This further reduces our capability of hydrological understanding and compromises forecasting in areas with emerging water scarcity, where these tools are needed most (Sivapalan, 2003). The low understanding we have of "ungauged basins" forces modellers and experimentalists to look beyond the classical approach of rainfall-runoff "curve-fitting" and to come up with either robust alternative strategies to find behavioural parameter distributions or to use other data sources than classical streamflow series. Kuczera (1983) showed that a strategy where information (not necessarily streamflow) can jointly be used to reduce uncertainty is updating of prior likelihoods by means of Bayes' law.

$$p\left(\Theta_i|Y,M\right)=\frac{p\left(Y|\Theta_i,M\right)p\left(\Theta_i|M\right)}{p(Y)} \qquad (1)$$

with

$$p(Y)=\sum_{i=1}^{\infty}p(Y|\Theta_i,M)p(\Theta_i,M) \qquad (2)$$

where the left-hand side represents the posterior probability of parameter set Θ_i of a given model M and the right hand-side represents the process of Bayesian updating with observation Y and joint prior distribution $p(Y)$. More and more modellers are applying this or similar methods, some for the purpose of taking into account new stream flow data as time proceeds (e.g. Freer et al., 1996), some for the purpose of presenting uncertainty based on a joint posterior parameter distribution (Kuczera, 1983) and some for the purpose of learning and consequently detecting model structural deficiencies and henceforth improving the model structure (Vaché and McDonnell, 2006; Son and Sivapalan, 2007; Fenicia et al., 2008).

In the era of remote sensing, new potentially interesting data sources emerge that may allow us to step-wise infer constraints on parameter distributions based on space-based measurements. Even though these data sources are often subject to a great deal of noise, resulting in a substantial and ill-quantifiable uncertainty, they can be employed as "soft data" (Seibert and McDonnell, 2002) to update prior likelihoods of parameters by Bayesian updating or to impose a certain degree of acceptance of a parameter set. Franks et al. (1998) made an attempt to constrain a model's parameter space by Bayesian updating, first using streamflow and consequently Synthetic Aperture Radar (SAR) estimates of saturated areas. In their study, a fuzzy estimate rather than the estimate itself, of the total saturated area, derived from SAR, was used as a way to deal with uncertainties related to subjective choices within the estimation procedure and the observations used. Franks and Beven (1997) included fuzzy estimates of Landsat TM derived evaporation estimates in the uncertainty reduction of a land surface model.

Other examples where remotely sensed information is used for calibration (rather than Bayesian updating) are described by Campo et al. (2006), Johrar (2002) and Immerzeel and Droogers (2008). In the former, a correlation of SAR backscatter with the top-soil moisture content of a distributed hydrological model of the Arno ($\sim$8230 (km)^2) was assumed. The authors recognised the problems of SAR soil moisture inference over vegetated areas. Moreover, SAR observations are difficult to apply on the large scale. In the latter two, first attempts of calibration with space-based evaporation were performed. Johrar (2002) calibrated an agro-hydrological model on remotely sensed evaporation rates and validated with in-situ measurements of groundwater levels. Immerzeel and Droogers (2008) calibrated a SWAT model for the Bhima catchment ($\sim$45 000 (km)^2) on SEBAL (Bastiaanssen et al., 1998) evaporation estimates by using a global optimisation algorithm. The Bhima catchment does not generate any streamflow at all, which means that for calibration, one has to rely on other information. A marked difference between the first and the last two studies is that actual evaporation is a flux, completely equivalent to evaporation from a hydrological model, while SAR 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. Liu and Gupta, 2007).

In this paper, an attempt is made to transfer a prior distribution of parameters that relate transpiration to soil moisture states, into a justifiable posterior distribution, by constraining the priors on satellite-based evaporation estimates during a dry season in the Luangwa basin in Zambia. The thermal-infrared character of SEBAL compromises its application in the cloud covered wet season. In this study Y in Eq. (1) is the space-based evaporation and M is given. We emphasize that the parameters that determine the depletion of the soil moisture zone in hydrological models, also determine to a large extent the separation of rainfall into soil moisture and streamflow. If the approach followed is successful, then this opens up new opportunities for constraining rainfall-runoff models in ungauged basins.

The analysis is performed on a semi-distributed conceptual model of the Luangwa river basin, a tropical area in Zambia, where recent information on stream flow is not available. At the moment, the basin is clearly ungauged, having no reliable stream-flow records after 1980 and poorly concomitant available time series of streamflow and rainfall (Fig. 1). The hypothesis is that the evaporation estimates impose a two-fold constraint on model parameters. First, the model is forced to obey the water balance and a realistic depletion of soil moisture within the dry season. According to old monthly records of rainfall and streamflow, evaporation accounts for about 85% of the annually averaged water balance in this river basin so it is a strong prescriptor for the water balance. Second, the modeller can attempt to regionalize evaporation sensitive parameters making use of observed land cover. An additional benefit of this approach is that the accuracy of the rainfall estimates that is used to force these models, is not of direct importance for this step in the reduction of parameter uncertainty, as long as the moisture status in the end of the wet season is more or less accurate.

Evaporation estimates based on thermal-infrared satellite imagery, are based on indirect estimation procedures, which may introduce a great deal of ill-quantifiable uncertainties (Wagener and Gupta, 2005). These are caused for instance by transferring 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) the somewhat subjective choice of a “wet” and “dry” pixel as extremes in the surface energy balance (described in a later section). Therefore, the evaporation estimates obtained, are in this study used as a proxy for the response of the land surface, assuming that errors are uncorrelated in time and space and are of a random nature.

2 Study area description

The Luangwa basin (Fig. 2) is a unique study area. It is a relatively pristine and remote area with only a small amount of old hydrometric station data (at the time of writing only one was fully operational, installed in November 2007). The Northern, most upstream part of the basin is mountainous and is subject to many locally generated flash floods. More downstream parts consist of sandy/loam soils (among which black cotton soils) covered by typical tropical savanna vegetation such as *Miombo* and *Mopane* forests (Frost, 1996). Many of these lower areas are interspersed with wetlands, locally called "dambos". The North-Eastern boundary (the Muchinga escarpment) consists of moist highlands covered by dense forested wetland areas, having a different hydroclimatology from the low lying savannas. Temperatures in the highlands are much lower and given the type of vegetation present, these areas have a higher capability of retaining moisture during the dry season than the lower savannah regions. The annual rainfall in the catchment is around 1000 mm per year (Fig. 2, right side). The heterogeneity of this area makes it an excellent site for research on the applicability of spatially distributed remotely sensed data in hydrological models.

The hydrological response of this area is crucial to the operation of the Cahora Bassa reservoir in Mozambique, the downstream riparian country. The Luangwa can generate critical and unexpected peak flows during Zambezi floods. The Luangwa joins the Zambezi, closely upstream lake Cahora Bassa. Operators of Cahora Bassa are sometimes forced to release large amounts of water from the reservoir, not knowing the exact magnitude of the Luangwa floods. In the past, this has caused unnecessary floods downstream of Cahora Bassa (e.g. in February and March 2001, February 2007), which resulted in loss of life and displacement of people and livestock from the

HESSD

5, 2293–2318, 2008

RR parameter constraints on evaporation

H. C. Winsemius et. al.

Zambezi flood plains. The Luangwa also carries a great deal of sediments. Siltation problems in Cahora Bassa are the result. The large carrying capacity of sediments has never been studied in detail as far as the authors are aware.

3 Material and methods

3.1 Evaporation estimates

To derive spatial-temporal variable estimates of the evaporation, the Surface Energy Balance Algorithm for Land (SEBAL) has been applied. An elaborate description of the current state of SEBAL is given by Allen et al. (2007). It was applied in several studies with varying applications described in 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. Unfortunately, evaporation cannot be assessed 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:

$$\rho_w \lambda E_a = R_n - G - H \quad (3)$$

where ρ_w is the density of water [$\mathrm{M\,L^{-3}}$], λ is the latent heat of vaporisation [$\mathrm{L^2\,T^{-2}}$] E_a is actual evaporation [$\mathrm{L\,T^{-1}}$], R_n is net radiation, G is ground heat flux and H is sensible heat flux [$\mathrm{M\,T^{-3}}$]. R_n is estimated from the satellite image derived broadband albedo, Normalized Difference Vegetation Index and surface temperature, together with incoming shortwave radiation estimates from Meteosat Second Generation (LSA SAF, 2007). G is estimated as a fraction of R_n, being dependent on the vegetation index. H is determined as an iterative solution to the equation

$$H = \rho_a c_p \left(T_s - T_{a,ref}\right) g_{ah} \quad (4)$$

HESSD

5, 2293–2318, 2008

RR parameter constraints on evaporation

H. C. Winsemius et. al.

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 "dry" and "wet" extremes in the satellite image). Then, The following assumption is made:

$$T_s - T_{a,ref} = aT_s + b \tag{5}$$

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, λE is found as the residual of Eq. (3). 24-h 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):

$$\rho_w \lambda \int_{t=0}^{t=24} \frac{E\,(t)}{R_n\,(t) - G\,(t)} dt = \frac{\rho_w \lambda E_{inst}}{R_{n,inst} - G_{inst}} \tag{6}$$

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).

$$g_s = \left(\frac{s\,(R_n - G) + \rho_a c_p\,(e_s - e_a)\,g_a}{\gamma \rho_w \lambda E} - \frac{s}{\gamma} - 1 \right)^{-1} g_{ah} \tag{7}$$

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, was taken from downscaled ECMWF fields. A physically based approach was followed to downscale the coarse grids of ECMWF to near-surface 1×1 (km)^2 variables (Voogt, 2006). A Jarvis model (Jarvis, 1976) was 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 was 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.

3.2 Conceptual model

A semi-distributed conceptual model has been set up for this study. Its distribution has been based on a one-year time series of decadal SPOT NDVI images. An unsupervised classification has been performed to differentiate between different land covers. The characteristics of the main dominant land covers was roughly defined through a short field investigation, elevation differences and, where available through investigation of high resolution, google earth overflight information http://earth.google.com/. Based on this information, 4 dominant land covers were defined: riverine, dambos (or wetlands), forested and highlands. The regions were manually delineated into polygon-shaped model units to decrease the computation time (Fig. 3, left side). The model structure applied, is a simplified version of the 1-dimensional box-model HBV (Fig. 3, right side, Lindström et al., 1997). The goal of this study is in first principle to investigate what the information density of the evaporation data is with respect to model identification, not to find the optimal model structure or parameter set for a given catchment. Therefore the amount of parameters was kept as low as possible in order to gain parameter identifiability. The model now consists of an interception store with a parameter D [L T^{-1}], an unsaturated soil zone S_u [L] (completely equal to the HBV soil zone), consisting of 3 parameters – in this paper referred to as S_{max} [L], B [–] and l_p [–], equivalent to the abbreviations "FC", "BETA" and "FLP" in the publication by Lindström et al. (1997). Outgoing fluxes are transpiration T_a [L T^{-1}] and recharge r_c [–], computed as

$$T_a = \min\left(\frac{S_u}{S_{max} l_p}, 1\right) T_p \tag{8}$$

$$r_c = \left(\frac{S_u}{S_{max}}\right)^B \tag{9}$$

Recharge is transferred to an upper zone S_q. Streamflow is generated from this zone, assuming it behaves as a linear reservoir with residence time K_q [T]. S_q represents the fast flow generated from water bodies or dambos. Finally, a lower zone S_s [L] (conceptualising groundwater) receives a maximum amount of percolation per time step from the upper zone, determined by the parameter F_{perc} [L T^{-1}]. This zone also behaves as a linear reservoir, contributing to the base flow with one parameter K_s [T] representing the average residence time. Note that in this study, the focus is only on the soil reservoir of the model, not on the runoff generating reservoirs. An upward flux from the runoff generating reservoirs to the soil reservoir was deliberately excluded. This results in a model structure in which there are only 2 parameters, S_{max} and l_p and one state, S_u, that influence the transpiration when there is no rainfall. Within the dry season, there is no significant sensitivity for parameter D and B, since they do not influence the depletion. D has been fixed on 2 mm day^{-1}. The prior value of the parameter B has been constrained, by making it dependent on soil texture. A normalised soil texture map was derived from the WISE-ISRIC dataset (Batjes, 2006), by weighting for each soil class the different percentages of present soil types with their respective texture class (coarse (0), medium (0.5) or fine (1)). B was roughly estimated by assuming that it has a value ranging between 1 and 4.5, linearly depending on the normalised soil texture between 0 and 1. Also the routing parameters F_{perc}, K_s and K_q do not influence our results, because there is no feedback from the discharge generating reservoir towards the soil moisture. Therefore these parameters are not mentioned in the remainder of this article.

3.3 Parameter distribution

The next step involves estimation of the two remaining evaporation-sensitive parameters S_{max} and I_p using the SEBAL estimates as evaluation data set. Parameter sampling and model evaluation was done separately for each model unit. To this end, SEBAL evaporation estimates have been lumped per model unit. 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 appears to be cloud-covered (and thus excluded from the SEBAL computation), the value has been discarded for the evaluation. The evaporation, generated from 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 at least unbiased. Uniform prior distributions were imposed for both S_{max} (varying between 200 and 2000 mm) and I_p (between 0.3 and 0.95). A model run time from September 2005 (driest moment in the year) until October 2006 was used. This period covers one rainy season, which allows a spin-up of the soil zone for the calibration period, May 2006 until October 2006. The following objective function L was used:

$$L(\Theta_i) = \frac{\sum_{t_p=1}^{m} \left[E_o\left(t_p\right) - E_s\left(t_p\right)\right]^2}{\sum_{t_p=1}^{m} \left[E_o\left(t_p\right) - \bar{E}_o\right]^2} \tag{10}$$

where t_p is a time period [T], m is the number of observation time periods available and E_o and E_s are the observed and simulated evaporation respectively [L T^{-1}]. L becomes lower as the model performance becomes better.

4 Discussion of the results

4.1 Model diagnostics and parameter identifiability

Of each model unit for which evaporation estimates of satisfactory quality were found, the best 2% of the parameter realizations (consisting of parameters S_{max} and I_p) was rescaled to a posterior likelihood L_s by

$$L_s(\Theta_i)=-\frac{L\left(\Theta_i\right)}{\sum_{j=1}^{n} L\left(\Theta_j\right)}+\frac{2}{n} \tag{11}$$

where n is the number of realizations, belonging to the best 2%. All in all, the sum of all the rescaled likelihoods L_s is equal to unity, where the highest values for L_s represent the best performing models. The results from all model units that have the same land cover class were binned together and plotted in Fig. 4. The number of model units belonging to the same class is indicated on the right-side of each sub-figure. For each land cover class, an example of a well-performing model is given in Fig. 5.

S_{max} is clearly clustered for all land cover classes. High optimal values are found in forested and highland regions, while riverine areas and dambos show relatively low values for the optimum of S_{max}. Forested model units show least clustering, although the parameter ranges suggest that S_{max} should at least be higher than 1000 mm. For 2 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 I_p.

4.2 Physical interpretation of parameter posteriors and validity of the model structure

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 and shrubs dominate these areas and it is well known that the grass wilts very soon after the rains have passed. This explains the rather persistent optimal value of l_p for all 3 dambo model units being close to 1. It implies that transpiration rates decline immediately when $S_u<S_{max}$. A relatively low optimal value for S_{max} is found which also concurs with the small root depth of the dambo-type vegetation. Riverine areas show more or less a similar pattern, although the amount of model units used to sample from the parameter space was limited to only 2 and the parameter clustering is far less pronounced.

The highlands on the other hand, are covered with evergreen forests that apparently have a large reservoir of water available for transpiration. l_p is rather low, indicating drought resistance, and S_{max} is high in these areas (2000 mm or even beyond the prior range). There is no optimum found for S_{max} and we 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. 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, is not suitable for these areas: first, the trees 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, shows a bi-modal distribution in the combined response surface of the four forested model units. This dual mode 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 *Mopane* (*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 improvement. 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.

4.3 Outlook

The results of this study show that, even with limited ground-truth knowledge, remotely sensed evaporation can condition both model structure and model parameters. When observing a natural river basin from above, patterns can be observed that are the result of evolution, which has resulted in a co-existence of ecosystems and hydrological behaviour. As a result, there is interdependence between vegetation, evaporation and runoff. This interdependence can clearly be identified in our model structures, which represent a simplified perception of nature, for instance in the parameters S_{max} and l_p in our conceptual representation of the soil moisture zone. These parameters influence both runoff and evaporation a great deal. It means that, although we condition the range of possible parameter realizations in a period without any forcing in terms of rainfall, it will have a great impact on parameter realizations in terms of discharge as well. The wide range of likely parameter realizations, found for forested regions, also reveals that we have not yet learned enough about the land surface to condition our models reliably on these data. Besides uncertainty in the evaporation data, there are still a lot of hypotheses on which we may condition our model structure and parameters (e.g.

difficulties in identification of the type of trees within forested model units, groundwater convergence in the highlands, missing knowledge about the geology, redistribution of runoff in wetlands). Nonetheless, it is far better to use these posterior distributions as a soft constraint (Seibert and McDonnell, 2002) than not to use them at all. Although evaporation estimates from space may be noisy and result in discontinuous response surfaces (as seen in Fig. 4), they offer one of the few, maybe even the only, opportunities to justifiably distribute parameters in ungauged basins. This would improve the decision making that can be done based on such hydrological models. Even as a soft constraint, for instance in the form of a fuzzy measure that may be employed as a first measure of degree of acceptance, these posteriors will impose a constraint for distributed land-surface related parameters in basins with little gauging. In this case, the red trapezoidal functions (see Fig. 4) could be used as fuzzy measures to act as prior in a subsequent calibration step on for instance (old) streamflow records. It is the view of the authors that when more and more of such soft constraints are included and combined with hard constraints in the modelling process, we will eventually be able to make better predictions in ungauged basins, including the necessary uncertainty assessments.

5 Conclusions

In this study, we presented a method with which remotely sensed data can be employed to construct posterior parameter distributions of land-surface related parameters in hydrological models. A simple conceptual representation of the soil moisture dynamics was chosen in order to present clearly identifiable parameter constraints. The results show that there is a clear consistency in the posterior likelihoods of parameters given different land cover classes. The consistent modes of the response surface are hydrologically meaningful in the sense that these modes 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. Because the used evaporation estimates are typically noisy, we feel that the posterior likelihoods found, can be used as a fuzzy rule to penalise likelihoods constructed from other information.

We argue that S_{max} may reach unrealistically high values, especially for the evergreen forest covered highlands, which may point in the direction of deficiencies in the model structure: the evergreen forests are probably capable of tapping from groundwater sources during the dry season, which may be replenished laterally. This is a process that is not yet included in the model structure and which could lead to better process understanding if included in the model structure.

Furthermore, the response surface for forested area exhibits a bi-modal distribution. This learns us that we probably have oversimplified the variability in forest types and their coverage. Multiple species of *Miombo* and *Mopane* forests co-exist in these areas where in general *Mopane* favours shallow and poor soils and *Miombo* has a deeper rooting system, dominating richer soils. This underlines the need for collaboration between hydrologists and ecologists for improving the understanding of the synergy between hydrology and ecology in large eco-systems and hence improve our spatial discretisation.

The approach followed, enabled us to spatially distribute and constrain some crucial parameters that determine the water balance in an ungauged basin of considerable size without any direct calibration on streamflow and, interestingly enough, without consideration of periods with direct forcing. Moreover, in semi-arid areas, where evaporation is a much larger water balance term than streamflow, this step should be preceding calibration on discharge in order to justifiably include spatial distribution of the water balance. The derived constraints will prove useful in cases of ungauged catchments, especially if used in a Bayesian updating framework, combining multiple constraints.

Acknowledgements. We are very grateful for the expert knowledge of, and cooperation and friendship with the staff of the Zambia Meteorological Department and the Department of Water Affairs of the Ministry of Energy and Water Development, Zambia. Furthermore, we are grateful for the cooperation with the Zambia Wildlife Authority, especially James Milanzi. A special thanks goes to our good friend Michael Mutale for all his logistic arrangements and provision of contacts. Finally, we are grateful for the financial support by the Netherlands Organisation for Scientific Research (NWO).

References

Allen, R. G., Tasumi, M., and Trezza, R.: Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model, J. Irrig. Drain. E.-ASCE, 133, 380–394, doi:10.1061/(ASCE)0733-9437(2007)133:4(380), 2007. 2300

Bastiaanssen, W. G. M., Hoekman, H. H., and Roebeling, R. A.: A methodology for assessment of surface resistance and soil water storage variability at mesoscale based on remote sensing measurements, in: IAHS Special Publication, 2, p. 65, IAHS Press, Wallingford, 1994. 2302

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., and Holtslag, A. A. M.: A remote sensing surface energy balance algorithm for land (SEBAL). Part 1. Formulation, J. Hydrol., 212–213, 198–212, 1998. 2297

Bastiaanssen, W. M. G., Ahmed, M. D., and Chemin, Y.: Satellite surveillance of evaporative depletion across the Indus Basin, Water Resour. Res., 38, 1273–1282, 2002. 2300

Batjes, N. H.: ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1.1), Tech. rep., ISRIC – World Soil Information, Wageningen, 2006. 2303

Beven, K. J. and Binley, A. M.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, 1992. 2295

Beven, K. J. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001. 2295

Campo, L., Caparrini, F., and Castelli, F.: Use of multi-platform, multi-temporal remote-sensing data for calibration of a distributed hydrological model: an application in the Arno basin, Italy, Hydrol. Process., 20, 2693–2712, 2006. 2297

Chidumayo, E. N.: Effects of climate on the growth of exotic and indigenous trees in central Zambia, J. Biogeogr., 32, 111–120, 2005. 2307

Farah, H. O.: Estimation of regional evaporation under different weather conditions from satellite and meteorological data, A case study in the Naivasha Basin, Kenya, Ph.D. thesis, Wageningen University, 2001. 2302

Fenicia, F., McDonnell, J. J., and Savenije, H. H. G.: Learning from model improvement: On the contribution of complementary data to process understanding, Water Resour. Res., 44, W06419, doi:10.1029/2007WR006386, 2008. 2296

Franks, S. W. and Beven, K. J.: Estimation of evapotranspiration at the landscape scale: a fuzzy disaggregation approach, Water Resour. Res., 33, 2929–2938, 1997. 2297

Franks, S. W., Gineste, P., Beven, K. J., and Merot, P.: 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, 1998. 2297

Freer, J., Beven, K. J., and Ambroise, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, Water Resour. Res., 32, 2161–2173, 1996. 2296

Frost, P.: The *Miombo* in Transition: Woodlands and Welfare in Africa, chap. The ecology of miombo woodlands, 11–58, Center for International Forestry Research, 1996. 2299

Fuller, D. O.: Canopy phenology of some mopane and miombo woodlands in eastern Zambia, Global Ecol. Biogeogr., 8, 199–209, 1999. 2307

Gragne, A. S., Uhlenbrook, S., Mohamed, Y., and Kebede, S.: Catchment modeling and model transferability in upper Blue Nile basin, lake Tana, Ethiopia, Hydrol. Earth Syst. Sci. Discuss., 5, 811–842, 2008, http://www.hydrol-earth-syst-sci-discuss.net/5/811/2008/. 2300

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeorol., 8, 38–55, 2007. 2315

Immerzeel, W. W. and Droogers, P.: Calibration of a distributed hydrological model based on satellite evapotranspiration, J. Hydrol., 349, 411–424, 2008. 2297, 2300

Jarvis, P. G.: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field, Philos. T. Roy. Soc. B., 273, 593–610, 1976. 2302

Johrar, R. K.: Estimation of effective soil hydraulic parameters for water management studies in semi-arid zones, PhD thesis, Wageningen University, 2002. 2297

Kuczera, G.: Improved parameter inference in catchment models, 2, Combining different kinds of hydrologic data and testing their compatibility, Water. Resour. Res., 19, 1163–1172, 1983. 2296

Lewis, D. M.: Observations of tree growth, woodland structure and elephant damage on Colophospermum mopane in Luangwa valley, Zambia, Afr. J. Ecol., 29, 207–221, 1991. 2307

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergstrom, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, 1997. 2302

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: Toward and integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007. 2297

LSA SAF: Validation report, Tech. Rep. SAF/LAND/IM/VR/1.6, LSA SAF, Lisboa, 2007. 2300

Mohamed, Y. A., Bastiaanssen, W. G. M., and Savenije, H. H. G.: Spatial variability of evaporation and moisture storage in the swamps of the upper Nile studied by remote sensing techniques, J. Hydrol., 289, 145–164, 2004. 2300

Monteith, J. L.: Evaporation and surface temperature, Q. J. Roy. Meteor. Soc., 107, 1–27, 1981. 2301

Penman, H. L.: Natural evaporation from open water, bare soil and grass, Proc. R. Soc. Lon. Ser.-A, 193, 120–145, 1948. 2301

Savenije, H. H. G.: Equifinality, a blessing in disguise?, Hydrol. Process., 15, 2835–2838, 2001. 2295

Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen, W. G. M., and Bierkens, M. F. P.: Assimilation of remotely sensed latent heat flux in a distributed hydrological model, Adv. Water Resour., 26, 151–159, doi:10.1016/S0309-1708(02)00089-1, 2003. 2300

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38, 1241, doi:10.1029/2001WR000978, 2002. 2296, 2308

Sivapalan, M.: Prediction in Ungauged Basins: a grand challenge for theoretical hydrology, Hydrol. Process., 17, 3163–3170, 2003. 2296

Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, Water Resour. Res., 43, W01415, doi:1029/2006WR005032, 2007. 2296

Uhlenbrook, S., Seibert, J., Leibundgut, C., and Rodhe, A.: Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, Hydrol. Sci., 44, 779–797, 1999. 2295

Vaché, K. B. and McDonnell, J. J.: A process-based rejectionist framework for evaluating catchment runoff model structure, Water Resour. Res., 42, W02409, doi:10.1029/2005WR004247, 2006. 2296

Voogt, M.: METEOLOOK, a physically based regional distribution model for measured meteorological variables, M.Sc. thesis, Delft, University of Technology, 2006. 2302

Wagener, T. and Gupta, H. V.: Model identification for hydrological forecasting under uncertainty, Stoch. Env. Res. Risk A., 19, 378–387, 2005. 2295, 2298

Young, P. C.: Model validation, chap. Data-based mechanistic modelling and validation of rainfall-flow processes, 117–161, Wiley, Chichester, 2001. 2295

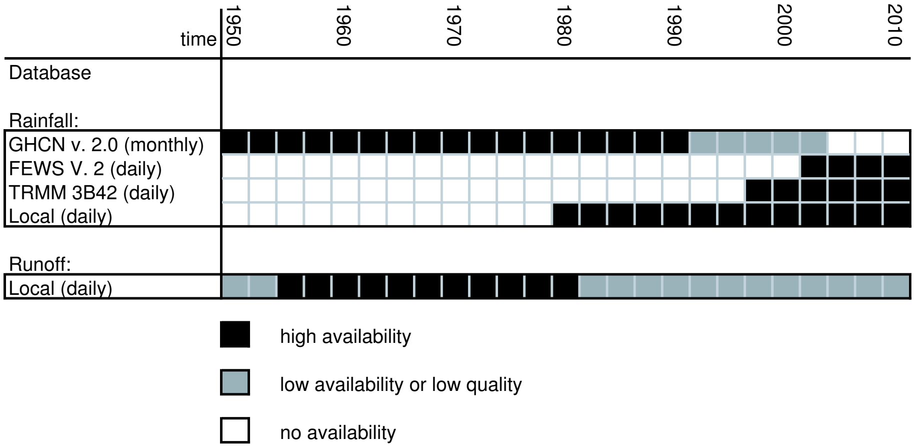

Fig. 1. Data availability in the Luangwa basin.

HESSD
5, 2293–2318, 2008

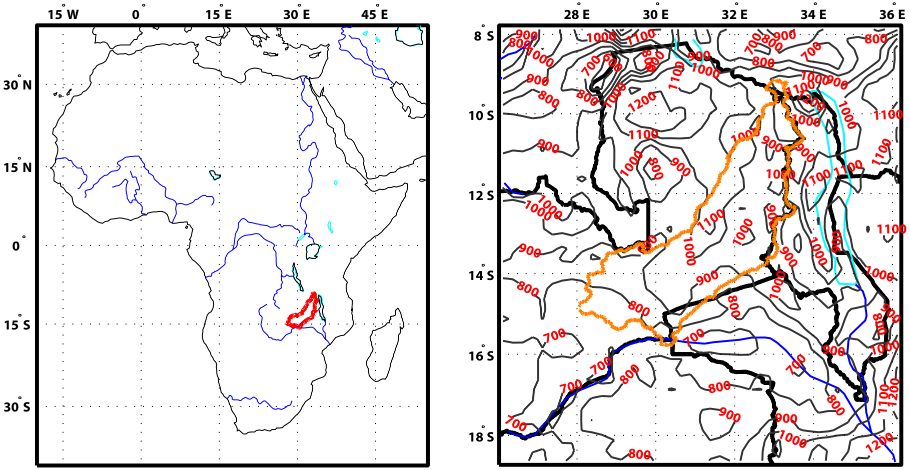

Fig. 2. Left: the study area, located in Southern Africa. Right: the study area (in orange) plotted on a terrain map, the isohyetes represent annual rainfall climatology, determined from the Tropical Rainfall Measuring Mission product 3B43 V6 (Huffman et al., 2007) averaged from 1998 until 2007.

HESSD

5, 2293–2318, 2008

RR parameter constraints on evaporation

H. C. Winsemius et. al.

HESSD

5, 2293–2318, 2008

RR parameter constraints on evaporation

H. C. Winsemius et. al.

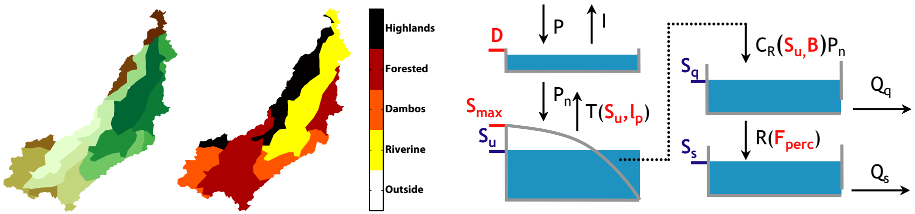

Fig. 3. Left: the delineated model units and dominant land cover. Right: the model structure.

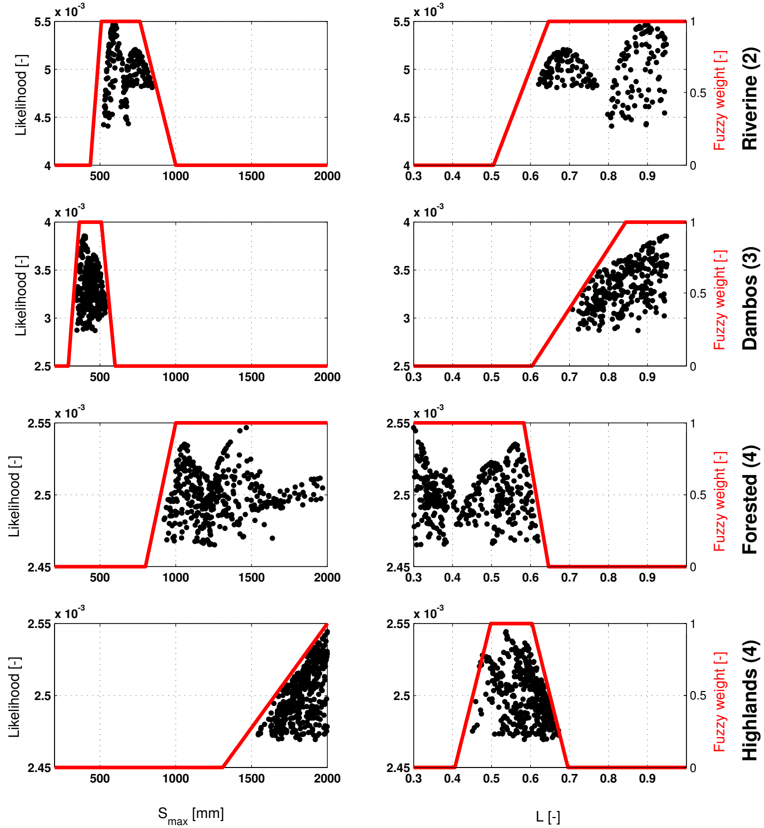

Fig. 4. Likelihoods of S_{max} and l_p given the SEBAL evaporation estimates. The results from different model units with similar land cover (the amount of model units is given on the right side between brackets) are combined in one figure to reveal resemblances in the response surface. The red lines indicate possible trapezoidal fuzzy measures that could be applied in a later calibration step as parameter constraint.

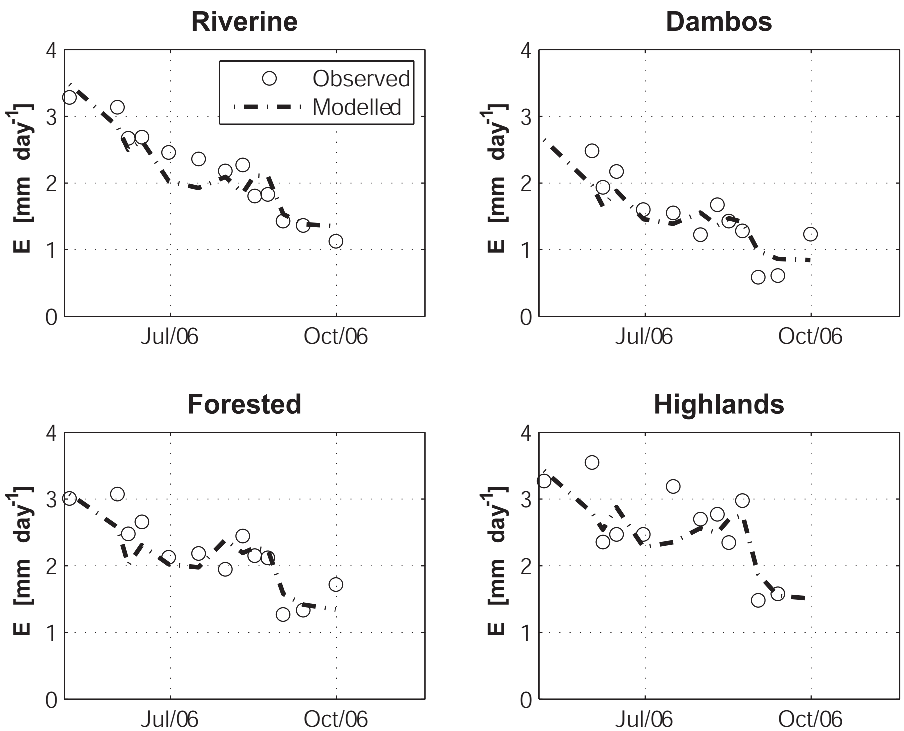

Fig. 5. Examples of the simulation performance of transpiration for well-performing parameter sets per land cover class.