From runoff to rainfall: inverse rainfall-runoff modelling in a high temporal resolution

M. Herrnegger[1], H.P. Nachtnebel[1] and K. Schulz[1]

[1] {Institute of Water Management, Hydrology and Hydraulic Engineering, University of Natural Resources and Life Sciences, Vienna, Austria}

Correspondence to: Mathew Herrnegger (mathew.herrnegger@boku.ac.at)

Abstract

Rainfall exhibits a large spatio-temporal variability, especially in complex alpine terrain. Additionally, the density of the monitoring network in mountainous regions is low and measurements are subjected to major errors, which lead to significant uncertainties in areal rainfall estimates. In contrast, the most reliable hydrological information available refers to runoff, which in the presented work is used as input for an inverted HBV-type rainfall-runoff model that is embedded in a root finding algorithm. For every time step a rainfall value is determined, which results in a simulated runoff value closely matching the observed runoff. The inverse model is applied and tested to the Schliefau and Krems catchments, situated in the northern Austrian Alpine foothills. The correlations between inferred rainfall and station observations in the proximity of the catchments are of similar magnitude compared to the correlations between station observations and independent INCA rainfall analysis provided by the Austrian Central Institute for Meteorology and Geodynamics (ZAMG). The cumulative precipitation sums also show similar dynamics. The application of the inverse model is a promising approach to obtain additional information on mean areal rainfall. This additional information is not solely limited to the simulated hourly data, but also includes the aggregated daily rainfall rates, which show a significant higher correlation to the observed values. Potential applications of the inverse model include gaining additional information on catchment rainfall for interpolation purposes, flood forecasting or the estimation of snow melt contribution. The application is limited to (smaller) catchments, which can be represented with a lumped model setup and to the estimation of liquid rainfall.

1 Introduction

The motivation for the concept presented in this paper comes from practical hydrological problems. Some years back we set up rainfall-runoff models for different alpine rivers (e.g. Stanzel et al., 2008; Nachtnebel et al., 2009a, 2009b, 2010a or 2010b). In the course of these projects, we were confronted with massive errors in the precipitation input fields. This is a known problem, especially in alpine environments. Although the temporal dynamics in the runoff simulations were captured quite well, significant mass balance errors between observed and simulated runoff were found. It could be excluded, that erroneous evapotranspiration calculations were biasing the results (Herrnegger et al., 2012). In the HYDROCAST project (Bica et al., 2011) we tested different precipitation interpolation and parameterisation schemes by using the ensemble of generated inputs for driving a rainfall-runoff model and comparing the simulated runoff time series with observations. In essence, the results showed, that no significant improvements could be made in the runoff simulations and that the information on the precipitation fields is strongly determined and limited by the available station time series. Runoff observations as an additional information source constitute a good proxy to precipitation observations with a considerably lower level of associated uncertainty. The main aim is therefore to present a proof-of-concept for the inversion of a conceptual rainfall-runoff model. That is to show, that it is possible to use a widely applied model concept to calculate mean areal rainfall from runoff observations.

Uncertainties in catchment precipitation

Areal or catchment precipitation estimates are fundamental, as they represent an essential input for modelling hydrological systems. They are however subject to manifold uncertainties, since it is not possible to observe the mean catchment rainfall itself (Sugawara, 1992; Valéry et al., 2009). Catchment rainfall values are therefore generally estimated by interpolation of point measurements, sometimes incorporating information on the spatial rainfall structure from remote sensing, e.g. radar (e.g. Haiden et al., 2011). Measurement, sample and model errors can be identified as sources of uncertainty. Point observations of rainfall, which are the basis for the calculation of mean areal rainfall values, are error inflicted (Sevruk, 1981, 1986; Goodison et al, 1998; Sevruk and Nespor, 1998; Seibert and Moren, 1999; Wood et al., 2000; Fekete et al., 2004). Occult precipitation forms like fog or dew are frequently ignored. Although not generally relevant, this precipitation form can be a significant contribution to the water budget of a catchment (Elias et al., 1993; Jacobs et al.,

2006; Klemm and Wrzesinsky, 2007). The highest systematic measurement errors of over 50 % are found during snowfall in strong wind conditions. Other sources of systematic measurement errors and their magnitudes are listed in Table 1.

➔ Approximate location of Tab. 1

In complex terrain the rainfall process is characterised by a high temporal and spatial variability. Especially in these areas the density of the measurement network tends to be low, not capturing the high variability and leading to sample errors (Wood et al., 2000; Simoni et al., 2011; de Jong et al., 2002). Further uncertainties arise in the interpolation of catchment scale rainfall from point observations. Systematic and stochastic model errors of the regionalisation methods can be identified. Systematic model errors can arise during the regionalisation of rainfall in alpine areas, when e.g. the elevation dependency is not considered (Haiden and Pistotnik, 2009). Quantitative areal rainfall estimates from radar products are, although they contain precious information on the rainfall structure, still afflicted with significant uncertainties (Krajewski et al., 2010; Krajewski and Smith, 2002). A general magnitude of overall uncertainty, which arises during the generation of areal rainfall fields, is difficult to assess, as different factors, e.g. topography, network density or regionalisation method, play a role.

Uncertainties in runoff observations

Errors in runoff measurements are far from negligible (Di Baldassarre and Montanari, 2009; McMillan et al., 2010; Pappenberger et al., 2006; Pelletier, 1987). When applying the rating-curve method for estimation of river discharge the uncertainties are a function of the quality of the rating curve and the water level measurements. The quality of the rating curve depends on (i) the quality and stability of the measured cross-section over time, (ii) the representativeness of the velocity measurements and (iii) the influence of steady and unsteady flow conditions. According to literature the overall uncertainty can vary in the range of 5% - 20% (Di Baldassarre and Montanari, 2009; Pelletier, 1987). Although it can be expected, that the measurement error will certainly be large during flood events due to its dynamic features, the errors are considerably lower compared to rainfall measurements and to the uncertainties introduced, when calculating mean areal rainfall. It must however be assumed, that transboundary flows and groundwater flows around the gauging station are negligible.

Catchment precipitation from runoff observations through inverse modelling

A classical application of hydrology, the problem of reproducing observed runoff with meteorological forcings as input through a formalised representation of reality, is a forward or direct problem. Two inverse problems related to this forward problem can be identified (Groetsch, 1993):

1. Causation problem: Determination of input I (=cause), with given output O (=effect) and given model K, including model parameters θ (=process)
2. Model identification problem: Determination of model K, given input I and output O

The model identification problem can be divided into (i) the problem of identifying the model structure itself and (ii) the determination of model parameters that characterise the system (Tarantola, 2005). The focus in this contribution lies in solving the causation problem, i.e. in the determination of rainfall input from runoff, with a given model structure and parameters. In the following, the model, which calculates mean catchment rainfall values from runoff, will be called *inverse model*. The conventional model, which uses rainfall and potential evapotranspiration as input to calculate runoff, will be called *forward model*.

Runoff from a closed catchment is the integral of rainfall minus evapotranspiration losses and change in water storage over a certain period of time. Therefore, runoff observations can be used to derive information on rainfall. This has been done in several studies, e.g. Bica et al., 2011; Valéry et al., 2009, 2010; Ahrens et al., 2003; Jasper and Kaufmann, 2003; Kunstmann and Stadler, 2005 or Jasper et al., 2002. The common basis of these studies was to indirectly gain information on catchment rainfall by comparing simulated runoff results with observations. Hino and Hasabe (1981) fitted an AR (autoregressive) model to daily runoff data, while assuming rainfall to be white noise. By inverting the AR model they directly generated time series of rainfall from runoff. Vrugt et al. (2008) and Kuczera et al. (2006) derived rainfall multipliers or correction factors from stream flow with the DREAM- and BATEA-methods, these methods however being computationally intensive. In a well-received study, Kirchner (2009) analytically inverted a single-equation rainfall-runoff model to directly infer time series of catchment rainfall values from runoff. The Kirchner model (when deriving the storage-discharge relationship directly from runoff data) only has a single parameter and does not explicitly need rainfall as driving input for calibration. Rainfall data is however needed for the determination of rainless periods for the estimation of the sensitivity function. Krier et al. (2012) applied the model of Kirchner (2009) to 24 small and mesoscale

125 catchments in Luxembourg to generate areal rainfall. No systematic differences in the quality
126 of the rainfall estimates are found between different catchment sizes. In periods with higher
127 soil moisture the rainfall simulations however show a higher performance. The parsimonious
128 approach of Kirchner (2009) is limited to catchments, where discharge is determined by the
129 volume of water in a single storage and which can be characterized as simple first-order
130 nonlinear dynamical systems. Also due to the larger number of model parameters describing
131 several linked storages, accounting for a variety of different runoff components, HBV-type
132 conceptual models offer higher degrees of freedom and flexibility in the calibration
133 procedure. They can, in consequence, be applied to catchments with a wider range of runoff
134 characteristics (Bergström, 1995; Kling et al., 2015; Kling, 2006; Perrin et al., 2001).
135 Therefore, in this study, the conceptual rainfall-runoff model COSERO (Nachtnebel et al.,
136 1993; Eder et al., 2005; Kling and Nachtnebel, 2009, Herrnegger et al., 2012; Kling et al.,
137 2015, among others), which in its structure is similar to the HBV-model, is used as a basis for
138 the inverse model. The COSERO model has been frequently applied in research studies, but
139 also engineering projects (see Kling et al., 2015 for details).

140 This paper is organized as follows: Following this introduction the methods section describes
141 the conventional conceptual rainfall-runoff model (forward model) and the inverse model,
142 including the preconditions and limitations of its application. The concept of virtual
143 experiments to test the invertibility of the inverse model and to analyse the effects of errors in
144 the discharge measurements on the inverse rainfall simulations are presented. Additionally,
145 the setup of different simulation experiments, e.g. to evaluate the influence of differing
146 calibration periods or possible runoff measurement errors on the simulations, are explained.
147 The inverse model is applied to two headwater catchments in the foothills of the northern
148 Austrian Alps, with differing hydro-climatic and physical conditions. The catchments and the
149 data base, including the calibration periods for the simulation experiments, is presented. The
150 runoff simulations of the forward model and the rainfall simulations of the inverse model are
151 described in detail in the results and discussion section. Finally the paper ends with a
152 summary and outlook.

2 Methods

2.1 Forward model (Rainfall-runoff model)

In the state space formulated forward model, the unknown runoff Q_t is a function f of known variables rainfall input R_t, potential evapotranspiration ETp_t, system states S_{t-1} and a set of model parameters θ_i, whereas the index t denotes time:

$$Q_t = f(R_t, ETp_t, S_{t-1} \mid \theta_i) \quad (1)$$

The rainfall-runoff model is based on the COSERO model (see introduction for references), but has a simpler model structure. It includes an interception and soil module and three reservoirs for interflow, base flow and routing. The model structure is shown in Fig. 1, model parameters are summarized in Table 2 and fluxes and system states in Table 3.

➔ Approximate location of Fig. 1

➔ Approximate location of Tab. 2

➔ Approximate location of Tab. 3

The COSERO-model is formulated in a state space approach, with state transition functions

$$S_t = f(S_{t-1}, I_t \mid \theta_i) \quad (2)$$

and output functions

$$O_t = g(S_{t-1}, I_t \mid \theta_i) \quad (3)$$

with

I_t Input, e.g. rainfall

O_t Output, e.g. total runoff

S_t System states, e.g. water stored in soil module

θ_i Model parameters.

So, the model state and the output at time t depend only and exclusively on the previous state S_{t-1}, the inputs I_t and parameters θ_i. The simplified model formulation can be found in the appendix.

2.2 Inverse model (Runoff-rainfall model)

In the inverse model the unknown rainfall R_t is a function of runoff Q_t, potential evapotranspiration ETp_t, system states S_{t-1} and a given set of model parameters θ_i, where again the index t denotes time:

$$\mathrm{R_t} = f^{-1}(Q_\mathrm{t}, ETp_\mathrm{t}, S_{\mathrm{t\text{-}1}} \mid \theta_\mathrm{i}) \tag{4}$$

If eq. (4) is invertible and given ETp_t, S_{t-1} and θ_i, there is only one single input I_t, which results in an output O_t (eq. (3)). To calculate the inverse rainfall rate the forward model is embedded in a search algorithm, to find, for every time step t, the rainfall rate R_t that best fits the observed runoff:

$$\mathrm{f(R_t)} = QSIM_\mathrm{t}(R_\mathrm{t}, ETp_\mathrm{t}, S_{\mathrm{t\text{-}1}} \mid \theta_\mathrm{i}) - QOBS_t \leq \varepsilon \tag{5}$$

with

$$\mathrm{R_{t,min}} \leq \mathrm{R_t} < \mathrm{R_{t,max}} \tag{6}$$

The upper and lower brackets of rainfall ($R_{t,min}$ and $R_{t,max}$) is set to 0 and 50 mm/h. The value of the upper bound is an arbitrary value, but any reasonable bounds can be applied. $QSIM_t$ and $QOBS_t$ are the simulated and observed runoff respectively. ε denotes a small value, which is ideally zero.

Solving eq. (5), which reflects the objective function used in the search algorithm, is basically a root finding problem. Different root finding algorithms were tested, with the Van Wijngaarden–Dekker–Brent Method (Brent, 1973; Press et al., 1992) being the method of choice, as this method exhibited the fastest results. The Brent method combines root bracketing, bisection and inverse quadratic interpolation to converge from the neighbourhood of a zero crossing and will always converge, as long as the function can be evaluated within the initial defined interval (in our case $R_{t,min}$ and $R_{t,max}$) known to contain a root (Press et al., 1992). The iteration progress for one model time step is illustrated in Fig. 2. The left y-axis shows the objective function values, the right y-axis (in logarithmic scale) the associated rainfall values estimated during the iteration procedure.

➔ Approximate location of Fig. 2

The state space approach of the model is a first order Markov process: The system states S_t and outputs O_t of the calculation time step depend only on the preceding states S_{t-1} and some

inputs I_t and not on the sequences of system states, that preceded it, e.g. S_{t-2}, S_{t-3}, ..., S_{t-n} (see eq. (2) and eq. (3)). All information of the sequence of the preceding inputs (I_{t-1}, I_{t-2}, ..., I_{t-n}) is implicitly included in the last relevant system state S_{t-1}. No hysteretic effects are considered in the model and it does not include a parameter, which introduces a lag effect between inputs and outputs.

Given the model structure, parameters and potential evapotranspiration as input, the inverse rainfall and resulting runoff are solely a function of the initial cold system states. The influence of the initial cold system states on the inverse rainfall calculation are analysed in the results section.

The determined rainfall value R_t represents the "best" simulated rainfall of the catchment and is also used as input into the forward model to simulate runoff. Therefore, for every time step the inverse model simulates a rainfall and corresponding runoff value and also resulting system states. The simulated runoff value should ideally be identical to the observed value. This is however not always the case, as will be shown later.

A more elegant method to calculate rainfall from runoff is by analytically inverting the equations of a given model, i.e. bringing the rainfall term onto the right side of the equation. In Herrnegger (2013) this method was presented, but showed some disadvantages. The model structure, which was used in Herrnegger (2013) and which can be inverted analytically, differs from the model presented here. It does not include interception and routing. Additionally the inversion is not possible in certain periods, since the discontinuities introduced by threshold values lead to non-invertibility in the analytical solution. The precondition that the rainfall-runoff model is invertible is violated in certain periods. For the forward model used here, an internal time step discretization is included in the model code to guarantee, that the transition between system states above and below the threshold value within a time step are solved exactly. This is not possible in the analytical solution presented in Herrnegger (2013), since no internal time step discretization can be implemented.

2.2.1 Preconditions and limitations of the application of the inverse model

It is assumed that runoff from the catchment passes through the measurement cross-section of the gauging station and that subsurface and transboundary flows are negligible. It is difficult to apply the inverse model to leaky catchments or catchments, where a significant part of the runoff is not observed at the gauging site. Even with a given quantification of the leakage

process, the application of the inverse model would lead to an additional uncertainty difficult to quantify. Since a novel approach is presented, it is also reasonable to exclude this possible source of error at this point. This is however not necessarily a limitation of the inverse model. Also the application of a forward hydrological model, which needs to be calibrated against runoff observations, will fail or will result in wrong estimates of water balance components.

The inverse model is based on a lumped model setup and the resulting inverse rainfall value corresponds to the mean areal rainfall. Applying a spatially distributed model is not possible, since the origin of outputs of different zones or cells of a distributed model setup cannot be reproduced by the inverse model in a deterministic way without additional assumptions. The information of origin gets lost as soon as cell values are summed and routed to a catchment runoff value. It is however conceivable to spatially disaggregate the mean areal rainfall from the inverse model using additional information, e.g. assuming an elevation dependency of rainfall.

Solid precipitation is accumulated without any direct signal on the hydrograph. It is therefore impossible to use the inverse model to estimate solid precipitation. The inverse model can only be used to calculate rainfall in snow-free catchments, or, as in our case, periods, in which runoff is not influenced by snow melt (i.e. summer months). However, in rainless periods, where it is clear, that snow melt is dominating runoff (e.g. in spring), the inverse model can be used to quantify snow melt rates from a catchment.

The applicability of the inverse model is limited to catchments, which are representable with a lumped model setup and the proposed model structure. If a catchment is too large, it will be generally difficult to simulate that system with a lumped model setup - not necessarily because of neglecting spatial heterogeneity in the model parameters (although this may also be an issue) or ignoring a lag between the rainfall and runoff signal, but simply because the lumped rainfall input used is “wrong” and is not representable for the whole catchment. If it only rains in the headwaters of large catchment, the lumped input into the forward model for this time step or rainfall event will be much lower, since it will be spatially aggregated. This input is not applicable to the whole catchment and the simulations will show deficits. In this case, an inversion will be highly flawed. This consideration is independent of the fact that the sampling of rainfall field in larger catchments tends to be statistically better, compared to smaller catchments, where observations are rarer.

It is also clear, that catchments, independent of size, exist, where the application of this particular model structure will fail (e.g. flatland catchments dominated by groundwater). If hydro-meteorological conditions of the catchment change or are different from the calibration period and the forward model (e.g. due to poor parameter estimation, inadequate model structure, wrong representation of the real world prototype etc.) is not able to capture these changes, then again the calculation of rainfall from runoff will fail (as they do for the forward case). However, being able to fit the forward model to observed runoff data and as long as the forward model is able to represent the catchment responses to rainfall, an inversion will be possible.

2.3 Simulation setups

2.3.1 Virtual experiments

In a first step the inverse model is evaluated and tested with virtual experiments, in order to guarantee, that the model equations are invertible. Runoff simulations are performed with the forward model driven by observed rainfall as input. The simulated runoff time series of the forward models are then used as input into the inverse model, with the aim to reproduce the observed rainfall. Simulated runoff from the forward model is dependent on the model parameters. Therefore, to test the inversion procedure for the whole parameter range, synthetic hydrographs are produced with Monte Carlo simulations. 20 000 different parameter combinations are chosen randomly from the parameter space, with the same number of model runs to evaluate the inverse model. The sampled parameters and associated range are shown in Table 2. The schematic setup of the virtual experiment and the evaluation of the inverse model is shown in Fig. 3. Note, that the setup and the evaluation is performed for every individual Monte Carlo run, as the simulated runoff from the forward model varies, depending on selected model parameters.

➔ Approximate location of Fig. 3

All system states and fluxes of the forward model are perfectly known at every time step. This information is used to evaluate the inverse model. Only after a successful evaluation of the inverse model with the virtual experiments, can observations of runoff be used as input into the inverse models.

Additionally, virtual experiments are performed, in which random noise drawn from a zero-mean normal distribution and rescaled to represent a range of measurement errors is added to a runoff simulation of the forward model. These time series are then used as input into the inverse model to test the sensitivity of the inferred precipitation rates to short-term errors in the discharge measurements:

$$Q\_FN_{i,t} = Q\_F_t + Q\_F_t * N\left(\mu, \sigma^2\right) * \alpha_i \quad (7)$$

with

$Q\_FN_{i,t}$ *Noisy input into inverse model*

Q_F_t *Forward simulated runoff based on observed precipitation*

$N(\mu, \sigma^2)$ *Normal distribution with mean* μ*=0 and standard deviation* σ^2 *=1*

α_i *Noise scaling factor: 0%, 1%, 2%, 5% and 10%*

2.3.2 Model calibration and simulations experiments with observed data

The application of the inverse model is based on the assumption that the forward model can represent the catchment responses to rainfall. The forward model is therefore calibrated against runoff observations, using observed rainfall values. The calibration setup and in consequence model parameters (for a given model structure) can depend on (i) the calibration period and length and (ii) the driving input used. The inverse rainfall is also a function of the observed runoff, which may also exhibit possible measurement errors. Finally, the initial conditions of the system states at the beginning of the simulations also influence the results of the forward, but also inverse model. To evaluate these influences, i.e. different model parameters due to different calibration periods and lengths, different runoff observations, different parameter optimisation data basis and different initial conditions, several simulation experiments are performed. An overview table of the simulation experiments can be found in section 3.3 (Table 5) after the presentation of the available data. Apart from the calibration period all simulation experiments include independent validation periods, which allow to test the inverse model in periods, in which no observed rainfall was used.

In a first step three different periods are used for calibration of the model parameters. In a further simulation experiment, the runoff observation is increased by a constant offset of 10% to evaluate the influence of possible systematic streamflow errors on the simulations and the

inverse rainfall. A fifth experiment is performed, in which an independent rainfall realisation is used as driving input for model calibration, in order to test the conditioning of the model parameters and in consequence the simulations to the driving input. Given the model structure, the inverse rainfall is a function of observed runoff, potential evapotranspiration, system states and model parameters (eq. (4)). Extending eq. (4) explicitly with all relevant system states leads to

$$\mathrm{R}_{\mathrm{t}} = f^{-1}(Q_{\mathrm{t}}, ETp_{\mathrm{t}}, BWI_{\mathrm{t-1}}, BW0_{\mathrm{t-1}}, BW2_{\mathrm{t-1}}, BW3_{\mathrm{t-1}}, BW4_{\mathrm{t-1}} \mid \theta_{\mathrm{i}}) \quad (8)$$

The forward and inverse models are run as a continuous simulation in time. The preceding system states are therefore an integral part of the simulation and are determined intrinsically within the simulation. However, the initial system states at the beginning of the simulation period (cold states) will influence the results of the simulation, but should, after an adequate spin-up time, not influence the runoff but also inverse rainfall simulations. Therefore, a sixth experiment was set up, in which three strongly differing cold start scenarios are defined:

- Reference scenario
- Dry system states scenario
- Wet system states scenario

For the reference scenario the system states from the continuous simulation were used. For the cold states in the dry scenario the states from the reference scenario were reduced by the factor 0.5 and increased by the factor 1.5 for the wet scenario.

The simulation experiments do not allow a systematic analysis of parameter uncertainty, since this is not the aim of this paper. The simulation experiments however enable a first assessment of the robustness of the results. That is to show the forward and inverse model performance, when the conditions are different from the conditions the model has been calibrated against (i.e. validation period) or if different driving inputs are used.

The model structure applied includes 12 parameters, of which 10 have to be calibrated. Two parameters (INTMAX and ETVEGCOR) are estimated a priori (see Table 2). The interception storage is represented by the model parameter INTMAX, which is estimated as a function of the land use and month of year to consider changes of interception within the annual cycle. ETVEGCOR, comparable to the widely used crop coefficient (Allen et al., 1998), is also estimated depending on the month of year and land use. Values for INTMAX

and ETVEGCOR can be found in Herrnegger et al. (2012). For the application, monthly INTMAX- and ETVEGCOR-values were calculated as area weighted mean values, depending on the land uses in the catchments, since a lumped setup is used. For the implementation of the evapotranspiration calculations in the model the reader is also referred to Kling et al. (2015).

Generally only June, July, August and September are used, since it can be guaranteed, that no snow melt influences runoff in these months (see section 2.2.1). Parameter calibration in the simulation experiments is performed for the forward model, using the Shuffled Complex Evolution Algorithm (Duan et al., 1992). As an optimisation criterion the widely used Nash-Sutcliffe-Efficiency (NSE, Nash and Sutcliffe, 1970) was chosen.

3 Materials

3.1 Study areas

The inverse model is applied to two catchments with different size, geology and land use located at the foothills of the Northern Alps. The Schliefau catchment is located about 110 km south-west of the Austrian capital of Vienna and covers an area of 17.9 km² with a mean elevation of 608 m.a.s.l.. About 55% of the area is covered by grassland and meadows, 40% by coniferous forest and 5% by mixed forest. The underlying geology is dominated by marl and sandstone. The Krems catchment is located about 170 km south-west of Vienna and covers an area of 38.4 km² with a mean elevation of 598 m.a.s.l.. The topography is more heterogeneous, with an elevation range of 413 to 1511 m.a.s.l., compared to 390 to 818 m.a.s.l. in the Schliefau catchment. Approximately 46% of the area is covered by grassland and meadows, 48 % by mixed forest, 4 % by settlements and 2 % by coniferous forest. On a long term basis, in both catchments, the highest runoff can be expected during snow melt in spring, the lowest runoff in summer and autumn until October. Fig. 4 shows a map of the catchments and Table 4 summarizes important characteristics of the study areas.

➔ Approximate location of Fig.4

➔ Approximate location of Tab.4

3.2 Meteorological database

Generally, two different rainfall time series are used. Ground observations of rainfall are available from the station St. Leonhard im Walde (Schliefau catchment) and Kirchdorf (Krems catchment), both located in the proximity of the catchments (Fig. 4). Additionally, areal rainfall data from the INCA system (Integrated Nowcasting through Comprehensive Analysis; Haiden et al., 2011) is used. INCA is the operational nowcasting and analysis application developed and run by the Central Institute for Meteorology and Geodynamics of Austria (ZAMG), which is also used for the majority of real-time flood forecasting systems in Austria (Stanzel et al., 2008). For the presented study analysis fields derived from observations, but no nowcasting fields, are used. Rainfall in INCA is determined by a nonlinear spatial interpolation of rain-gauge values, in which the radar field is used as a spatial structure function. In addition an elevation correction is applied (Haiden and Pistotnik, 2009). The stations used for the interpolation of the INCA-rainfall fields are shown as triangles in Fig. 4. Note, that the stations St. Leonhard im Walde and Kirchdorf are not included in the INCA analysis, since they are operated by a different institution. The rainfall fields from the INCA system cover the test basins in a spatial resolution of 1 km². From the spatial data set mean catchment rainfall values are obtained by calculating area-weighted means from the intersecting grid cells.

Potential evapotranspiration input is calculated with a temperature and potential radiation method (Hargreaves and Samani, 1982).

3.3 Simulation periods

Runoff and rainfall data is available for the period 2006 to 2009 in a temporal resolution of 60 minutes, which is also the modelling time step. The virtual experiments are performed for a period of 4.5 months (15.5.2006 – 30.09.2006) resulting in 3336 time steps being evaluated. As described in section 2.3.2 different model calibration and simulation experiments are performed. An overview of these experiments is given in Table 5.

➔ Approximate location of Tab.5

4 Results and discussions

4.1 Virtual experiments

In the virtual experiments it could be shown, that the invertibility of the model equations is given. Using all 20 000 simulated hydrographs from the Monte Carlo runs, where the parameters were varied stochastically, the observed rainfall time series could be identically reproduced by the inverse model. Apart from the rainfall also all fluxes and system states where identical in the forward and inverse model runs. The comprehensive results from these virtual experiments are documented in Herrnegger (2013).

For the second set of virtual experiments station data from the Schliefau catchment with model parameters of Exp3 (see Table 5) were used as driving input in the forward model and the resulting runoff simulation in succession as input into the inverse model. To these resulting runoff simulations, however, noise with different magnitudes was added beforehand.

Depending on the magnitude of noise added to the runoff input time series, the inferred precipitation rates differ from the observed values, as is shown in Table 6. Without any noise the observed rainfall is reproduced exactly. With increasing noise a deterioration of the model performance is evident. Temporal aggregation leads to an increase in the correlation values, since the resulting noise in the inferred precipitation rates are smoothed out. The mean observed precipitation rate for the evaluated period in these virtual experiments is 0.21 mm for hourly precipitation, 1.26 mm for the 6h-sums and 5.03 mm for the daily precipitation rates. Based on these values, the mean quantitative bias ranges between - 0.6% and -6.3% relative to the mean observed rainfall, depending on added noise scaling factor of 1% to 10%. The inferred precipitation totals are higher, compared to the observed values, since the noise also leads to a quantitative bias between the runoff simulation of the inverse model and the runoff used as input. From the results it is clear that the inferred precipitation rates are sensitive to potential short-term errors in discharge measurements. Especially for the case, in which the noise scaling factor was set to 10%, assuming large short-term errors, the inverse model is not able to reproduce the disturbed input time series. This is also evident from the mean squared error values. The noise with a scaling factor of 10% however leads to a strongly perturbed runoff time series. Also in the forward case it would not be able to reproduce this runoff time series with the given precipitation in a reasonable manner.

➔ Approximate location of Table 6

4.2 Forward model: Parameter calibration and validation of the different simulation experiments

A precondition for the application of the inverse model is that the observed runoff characteristics of the catchment are reproduced reasonably by the forward model, since these parameters are also used in the inverse model. The following section therefore presents the runoff simulations of the forward model, based on the different simulation experiments Exp1 to Exp5.

The model performance for different periods of the forward model, expressed by Nash-Sutcliffe-Efficiency (NSE) and the mean bias between simulated and observed runoff in percent of observed runoff is shown in Table 7. As mentioned before, only the months June, July, August and September of the single years are used.

➔ Approximate location of Tab.7

With the exception of Exp5 the NSE-values of the calibration periods are larger than 0.8 in both catchments. The highest NSE-values of 0.87 (Schliefau) and 0.88 (Krems) are found for Exp1. The short calibration period used in this experiment (only June to September 2006 are used; see Table 5) enables a good fitting of the model parameters to the runoff observations. In consequence the largest deterioration of the model performance in the validation period is evident for Exp1 for both catchment, since runoff conditions differ from calibration. For the other experiments the differences in the NSE-values between calibration and validation period are less pronounced, with some experiments showing higher model performance in the validation period. In Exp5 INCA rainfall data is used as driving input for the simulations. The main intention of Exp5 is to evaluate the influence of a different rainfall input on the calibration of the model parameters and in consequence also on the inverse rainfall. For both catchments, the NSE-values of the forward model are mostly significantly lower, also compared to Exp3, which has the same calibration and validation periods. Although INCA uses a complex interpolation scheme, also incorporating radar data and a rainfall intensity depending elevation correction (Haiden et al., 2011; Haiden and Pistotnik, 2009), it seems that the data set has deficits representing catchment rainfall compared to the station observations in the proximity of the catchments. This can be explained by the larger distance of about 10 to 35 km of the INCA stations from the catchment (see Fig. 4). Note, that the ground observations in the proximity of the catchments are not used in the interpolation

process for the INCA-rainfall fields, as they belong to a monitoring network operated by a different institution.

For Exp1 to Exp3, the NSE-values for the period 2006 to 2009 show, that the overall model performance is fairly stable and comparable, independent of the calibration length. The NSE-values are larger than 0.82, with the exception of Exp1 in the Krems catchment. Although the calibration lengths and periods in Exp2 and Exp3 differ, identical model parameters were found for the Krems catchment in the optimisation for both simulation experiments. As a consequence the model performance is identical in these two experiments for the period 2006 to 2009.

The mean bias does not show a clear pattern and seems to be independent from the calibration period and length. In the Schliefau catchment observed runoff is overestimated by 7.8 to 0.9 % and underestimated by -1.4 to -4.8% in the Krems catchment for the period 2006-2009, depending on the simulation Exp1 to Exp3. Overall the calculated bias between observed and simulated runoff is in reasonable bounds.

In Exp4 the observed runoff is increased by 10%, mainly to evaluate the influence of possible streamflow errors on the simulations and the inverse rainfall. The same calibration periods were used as in Exp3, with station observations as driving input into the model. The NSE of Exp4 is comparable to Exp1, Exp2 and Exp3. The mean bias in Exp4 however becomes larger in both catchments. The observed runoff is now also underestimated in the Schliefau catchment, what is not surprising, since observed runoff was increased.

Fig. 5 shows the NSE-values of the forward model for the calibration periods of every simulation experiment versus single years performance for the 2 study areas.

➔ Approximate location of Fig.5

For Exp1 a significant larger spread in the model performance within the single years is evident. In Exp1 only 2006 was used for calibration. As a consequence, especially for the Krems catchment, the model performance is lower in the years 2007 to 2009, compared to Exp2 and Exp3. In the short calibration period of 2006 the model parameters are overfitted to the observations. If the conditions in the catchment are different from the calibration period, the model performance can be expected to deteriorate, as has been shown before (e.g. Kling, 2015; Seibert, 2003) and explains the findings. For Exp2 to Exp4 the model performance is however stable for the single years, also for 2009, which was not used for calibration in any

simulation experiment. In contrary to the Krems area, a large spread in the model performance of the single years for Exp5 is visible in the Schliefau catchment. The reason is not clear and may be explained by changing availability of station data for the INCA rainfall in the single years. We can however not verify this hypothesis, since we do not have access to the data sets. In the Schliefau catchment low NSE-values are calculated for the year 2008 for all simulation experiments. In the beginning of June a flood was observed (Fig. 6), which is not simulated in the model runs and explains the lower NSE values in this year. Excluding this event in the performance calculations would result in a significantly higher NSE of 0.84 for Exp1 for the year 2008, compared to 0.63 when the flood event is included in the calculation.

Fig. 6 (Schliefau) and Fig. 7 (Krems) exemplarily show the runoff simulations based on the results of Exp2. For both catchments, the dynamics and variability of the runoff observations are mostly reproduced in a satisfactory manner. However, a tendency is visible, that larger floods are underestimated in the simulations.

➔ Approximate location of Fig.6

➔ Approximate location of Fig.7

All simulations are performed with a lumped model setup. Consequently heterogeneity in geology and land use within the catchment are not considered in the parameter estimation. Also taking this into consideration, it can be concluded that the general responses of the catchment to rainfall input are captured appropriately by the forward model. Only for Exp1 with the very short calibration period, a larger deterioration of the model performance in the validation period and a larger spread in independent years is evident. It is therefore justified to calculate areal rainfall from runoff using the inverted forward model, including the optimised parameters.

4.3 Inverse model

For the evaluation of the simulated rainfall from the inverse model (PInv) we will compare the calculated values with observed station data (PObs) of St. Leonhard (Schliefau catchment) and Kirchdorf (Krems catchment) and the rainfall values from the INCA-system (PInca). In the following cumulative rainfall sums and the correlation and bias between simulated and observed rainfall are presented. Additionally the rainfall and runoff simulations of a flood event and the influence of cold system states on the simulations are shown.

4.3.1 Cumulative rainfall sums

Fig. 8 and 9 show the cumulative curves of the observed rainfall (PObs), INCA rainfall (PInca) and the inverse rainfall (PInv) of the simulation experiments Exp1 to Exp5 for the Schliefau and Krems catchment. Additionally the cumulative observed runoff (Qobs) is shown as a dashed line. Note that for the Krems catchment (Fig. 9) the rainfall curves of Exp2 and Exp3 are identical, since the model parameters are also identical in these simulation experiments.

➔ Approximate location of Fig.8

➔ Approximate location of Fig.9

The cumulative sums of the inverse rainfall and the observation based rainfall realisations PObs and PInca mostly show very similar temporal dynamics. Although large deviations are sometimes evident for both catchments, the deviations of the cumulative curves of PInca and the different inverse rainfalls (PInv) from the cumulative curves of the ground observation (PObs) are mostly of similar magnitude.

The inverse rainfall curves of Exp1 to Exp5 of the two catchments do not exhibit substantial differences, although different calibration periods and setups were used. At the beginning of June 2008 a flood was observed in the Schliefau catchment, which was underestimated in the forward simulation, presumably due to inadequate representation of the storm event in the rainfall observations (see runoff simulation in Fig. 6, lower left). Larger rainfall intensities are therefore calculated by the inverse for this period, leading to the larger deviations between the cumulative sums of PObs and PInv of Exp1 to Exp5 as shown in Fig. 8 (lower left). In the Schliefau catchments larger differences between Exp1 to Exp5 occur in the year 2009 (Fig. 8, lower right). Here, in the second half of June, a period of strong rainfall is evident, which also led to a series of floods in the catchment (see also the hydrographs in Fig. 6). The rainfall sums originating from these high flows were calculated differently in the inverse models, depending on the simulation experiment. In consequence, the inverse rainfall curves differ from July onwards. In 2009, which was the wettest summer in both catchments, the highest inverse rainfall sums are found for Exp4. This is what could be expected, since the observed runoff was increased by 10% in this simulation experiment. However, in the other years Exp4 does not necessarily show the largest inverse rainfall sums. The optimised model parameters in Exp4, that control evapotranspiration, were limiting actual evapotranspiration from the

model to fulfil the water balance, since PObs was not changed. In the second half of June 2009, during the flood events with low evapotranspiration, the higher runoff values used as input however show a clearer signal in the inverse rainfall sums.

The large difference between cumulative rainfall and runoff curves highlight the importance of actual evapotranspiration (ETa) in the catchments. For the Schliefau catchment the mean observed rainfall for the summer months of 2006-2009 is 678 mm. 266 mm are observed in the mean for runoff. Neglecting storage effects, a mean actual evapotranspiration of 412 mm can be calculated from the water balance. Over 60 % of rainfall are therefore lost to evapotranspiration. The mean actual evapotranspiration from the inverse model, depending on the simulation experiment, ranges from 352 mm to 362 mm, and are lower compared to the ETa calculated from the water balance. In the Krems catchment a mean runoff of 334 mm and rainfall of 600 mm, resulting in an actual evapotranspiration of 266 mm, is calculated. Although lower compared to Schliefau, nearly 45 % of rainfall are here lost to the atmosphere. The mean actual evapotranspiration from the inverse model, again depending on the simulation experiment, range from 276 mm to 310 mm. If the model would not capture ETa adequately, the cumulative rainfall curves would not follow the observations so closely.

On the basis of the different cumulative rainfall sums it can be concluded, that on a longer temporal basis, the inverse model is capable of simulating the catchment rainfall from runoff observations. This is also the case for independent validation periods and years, which were not used in the calibration. The results from the different simulation experiments do not differ substantially and show close correspondence to the observed data, except for a single summer in the Schliefau catchment.

4.3.2 Correlation and bias between simulated and observed rainfall

The performance of the inverse model expressed by the correlation coefficient is used to measure the models ability to reproduce timing and shape of observed rainfall values. It is independent of a possible quantitative bias. In the introduction the difficulties involved in the quantitative measurement of rainfall were discussed. It can however be assumed that a qualitative measurement, e.g. if it rains or not, will be more reliable. Table 8 shows the correlation values between ground observations and the different inverse rainfall realisations (PObs – PInv) and ground observations and INCA rainfall (PObs – PInca) for different periods and temporal aggregation lengths.

➔ Approximate location of Tab.8

The highest correlation values between POobs and PInv for the 1h-sums and calibration period are found for Exp1 with 0.71 (Schliefau) and 0.62 (Krems). For the other experiments the correlation values in the calibration period are lower (0.51 to 0.57 in the Schliefau area and 0.44 to 0.49 in the Krems catchment). For the validation period the correlation between POobs and PInv deteriorates in Exp1. For the remaining experiments, however, the correlation in the validation period is mostly higher, compared to calibration. This agrees with the finding from the forward simulation results, since better model performance in the validation period of the forward model also leads to a higher correlation between POobs and PInv. For the temporally aggregated 24-h sums the correlation values generally increase for the calibration and validation periods.

For the period 2006 to 2009 and 1h-sums, the lowest correlation values between POobs and PInv are found for the simulation results of Exp1 in both catchments. The highest correlation values are found for Exp2 in the Schliefau catchment and Exp2 to Exp4 in the Krems catchment. This agrees with the performance of the forward model presented in section 4.2.. The correlation of the 1h-sums between POobs and PInv is rather weak. However, the correlation between POobs and PInv is higher for all simulation experiments and 1h-sums compared to the correlation between POobs and PInca. This is interesting, since PInca is based on station rainfall observations and PInv is indirectly derived from runoff through simulations. With temporal aggregation the correlation values generally increase significantly for all combinations. Small differences or timing errors in the 1h-sums are eliminated with temporal aggregation. This is also the case for the INCA data.

For Exp1 to Exp4, the model parameters used for the forward and inverse model were automatically calibrated using the ground observation POobs as input. It could therefore be concluded that the model parameters are conditioned by POobs and that in consequence the fairly good agreement between POobs and PInv originates from this conditioning. Based on this hypothesis, calibrating the model with INCA data should lead to a better agreement between the INCA data and the corresponding inverse rainfall and a deterioration of the correlation between station data and inverse rainfall. For Exp5, the forward model was therefore calibrated with INCA data and the resulting parameters set was then used to calculate the inverse rainfall. The correlation between PInca and PInv for Exp5 is however not higher, compared to the other simulation experiments and Exp3, which had the same

calibration period. This excludes that the parameters are conditioned (at least for the rainfall simulations) by the input used for calibration. The comparison of Exp3 and Exp5 is critical and shows, that the inverse model provides reasonable results in the case where the forward model is calibrated with rainfall data that are independent from the observed catchment rainfall: The forward model exhibits significantly lower NSE in Exp5 compared to Exp3, which is expected because the forward model is driven with the lower quality INCA rainfall in Exp5 (see Tab. 7). The correlation between PObs and PInv however suggests that Exp5 is comparably representative of the rainfall dynamics as Exp3.

The correlations between PInca and PInv are generally very weak, with values ranging from 0.25 to 0.29 for the Schliefau and 0.39 to 0.445 for the Krems catchment.

For the period 2006 to 2009, the correlation between PObs and PInv for the 1-h sums ranges between 0.48 and 0.55, but is higher, compared to the correlation between PObs and PInca. In contrast Kirchner (2009) shows correlation values between simulated and observed rainfall of 0.81 and 0.88 for his two sites. The Schliefau and Krems catchments differ substantially in size, hydrological characteristics, land use or geology. The NSE values of the runoff simulations in Kirchner (2009) are higher, compared to the values presented here for the forward model. As a consequence the better performance in the rainfall simulations may be explained with the fact, that the Kirchner (2009) model better reflects the catchment conditions leading to runoff.

For the 24-h sums and the period 2006 to 2009 we calculate a correlation of 0.87 to 0.92, depending on the catchment and simulation experiment. Here Kirchner (2009) shows correlation of 0.96 and 0.97. Krier et al. (2012) present correlations between simulated and observed rainfall of 0.81 to 0.98, with a mean value of 0.91 for a total of 24 catchments, however only on the basis of data of a single year. The correlation in our results is therefore in the range of other studies. Unfortunately Krier et al. (2012) do not present NSE-values of the runoff simulations. It is therefore not possible to check the link between the performance of the forward model and rainfall simulations in their study.

Fig. 10 shows the correlation between PObs and PInv for the calibration periods of the simulation experiments Exp1 to Exp5 versus the correlation in single years for the two study areas. For the Schliefau catchment the largest spread in the correlation values of the single years is found for Exp1, which also corresponds to the performance of the runoff simulations of the forward model. For Exp2 to Exp5 a spread is also visible between the single years, but

differences are smaller. For the years 2006, 2008 and 2009 the correlation values in the Krems catchment do not differ substantially. Here however the correlation for the year 2007 is very low, independent of the simulation experiment. This may be explained by the comparatively dry summer of 2007. Also in the Schliefau catchment the correlation values are mostly lower in 2007, compared to the other years.

➔ Approximate location of Fig.10

Tab. 9 summarizes the mean bias in mmh^{-1} and mmd^{-1} between different rainfall realisations, evaluated for different periods and for 1h- and 24-h-sums. Except for Exp1 the bias is larger in the validation compared to the calibration periods.

For the period 2006 to 2009 and the Schliefau catchment, the bias between PInv and PObs is mostly higher, compared to the bias between PInca and PObs. Only Exp2, with a mean bias of 0.07 mmd^{-1}, is comparable to the bias between PInca and PObs of 0.02 mmd^{-1}. Exp2 also showed the highest performance in the runoff simulations concerning the NSE. In contrary, for the Krems catchment, the bias is lower between PInv and PObs for Exp1 to Exp3, compared to PInca and PObs. For Exp1 to Exp3 and the period 2006-2009 a mean bias of 0.14 mmd^{-1} (Schliefau) and 0.36 mmd^{-1} (Krems) is calculated. As a comparison, Krier et al. (2014) published mean bias values between simulated and observed rainfall of -3.3 to 1.5 mmd^{-1} (mean -0.35 mmd^{-1}) for 24 catchments on the basis of a single year. From all simulation experiments, Exp4 shows the largest bias, which is explained by the fact, that runoff was increased in this experiment. Here the increased runoff clearly shows a signal in the inverse rainfall.

➔ Approximate location of Tab.9

4.3.3 Rainfall and runoff simulations for a flood event

Fig. 11 exemplarily illustrates the temporal development of the different rainfall realisations and runoff simulations for the highest flood event in the Krems catchment. Results from Exp3 are shown. Compared to PObs and PInca the inverse rainfall PInv exhibits higher variability and higher intensities. The higher variability and oscillating nature of the inverse rainfall is explainable with the reaction of the inverse model to small fluctuations in runoff observations: In case of rising runoff observations, rainfall will be estimated by the inverse model. If the observed runoff decreases and the simulated runoff of the inverse model is larger than observed runoff, no inverse rainfall will be calculated, leading to the visible

692 oscillations. Fig. 11 (b) shows, that the forward model, driven with PObs as input,
693 underestimates both flood peaks. The forward model, driven with the inverse rainfall,
694 simulates the driven periods very well (Inverse QSim). However, especially the falling limb
695 after the second flood peak on the 07.09.2007 is overestimated by the inverse model. In this
696 period it is also visible, that in consequence no rainfall is calculated by the inverse model,
697 since simulated runoff is higher than observed runoff.
698 ➔ Approximate location of Fig.11

699 For a given time interval, the inverse model will yield an exact agreement between observed
700 and simulated runoff, as long as there is a positive rainfall value R_t to solve eq. (5). This will
701 be the case in periods of rising limbs of observed runoff (driven periods), as a rainfall value
702 can be estimated, which raises the simulated runoff value to match observation. On the
703 contrary, in periods of observed falling limbs (non-driven periods) the simulated runoff will
704 solely be a function of the model structure, its parameters and the antecedent system states, as
705 negative rainfall values are ruled out beforehand. This explains, why in periods, in which the
706 simulated runoff is higher than the observed value, no rainfall is calculated by the inverse
707 model.

708 4.3.4 Influence of cold system states on the inverse rainfall (Exp6)

709 To test the influence of cold states on the inverse rainfall simulations the simulation
710 experiment Exp6 was performed. Three different cold states (Reference, dry and wet system
711 states) were thereby defined (see section 2.3.2). Fig. 12 exemplarily shows the results of Exp6
712 for the Krems catchment.
713 ➔ Approximate location of Fig.12

714 From the monthly rainfall sums of the different model runs it is evident, that the inverse
715 rainfall calculations differ significantly at the beginning of the simulation. In the first month
716 the reference scenario results in a monthly rainfall sum of 30 mm, the dry scenario in 111 mm
717 and the wet scenario in only 9 mm. Generally the model will always strive towards an
718 equilibrium in its system states, which are a function of the model structure and parameters.
719 In the scenario “wet” a lot of water is stored in the states of the model at the beginning, with
720 the result, that little inverse rainfall is calculated. In the dry scenario on the other hand a
721 higher amount of rainfall is estimated, since less water is stored in the states at the beginning.

With time, however, the different system states converge. In consequence also the inverse rainfall values converge and after 9 months no differences are visible.

Extreme assumptions were made concerning the dry and wet scenarios, since the intention of Exp6 is to evaluate the general influences of the cold states and spin-up time on the inferred rainfall. Especially the long memory of the ground water storage explains the long warm-up period in the presented results. In practice reasonable cold states must therefore be defined at start-up, as is the case for forward models formulated in a state-space approach. After an adequate spin-up time the system states will however converge, leading to deterministic and unique inverse rainfall estimates.

5 Summary and outlook

A calibrated rainfall-runoff model (forward model) reflects the catchment processes leading to runoff generation. Thus, inverting the model, i.e. calculating rainfall from runoff, yields the temporally disintegrated rainfall. In this paper we applied a conceptual rainfall-runoff model, which is inverted in an iterative approach, to simulate catchment rainfall from observed runoff. The precondition of invertibility of the model equations is successfully tested with virtual experiments, in which simulated runoff time series are used as input into the inverse model to derive rainfall. Additional virtual experiments are performed, in which noise is added to the runoff input time series to analyse the effects of possible short-term errors in runoff on the inferred precipitation rates.

The approach is applied and tested in two study areas in Austria. The estimated inverse rainfall is compared with two different rainfall realisations: Apart of ground observations, areal rainfall fields of the INCA-system are used. Hourly data is available for the years 2006 to 2009. Only the months of June to September are used, as the inverse model can only be applied to simulate rainfall in periods, in which runoff is not influenced by snow melt (i.e. summer months).

In a first step, the forward model is calibrated against runoff observations. To evaluate the influences of (i) different model parameters due to different calibration periods and lengths, (ii) different runoff observations and (iii) different parameter optimisation data basis on the runoff and rainfall calculations, several simulation experiments are performed. Additionally the influence of different initial conditions on the rainfall simulations are evaluated.

752 The forward model mostly shows stable results in both catchments and reproduces the
753 dynamics and variability of the catchment responses to rainfall in a satisfactory manner. Only
754 the simulation experiment, in which a single summer was used for parameter calibration,
755 shows a larger deterioration of the model performance in validation period and independent
756 years. The model parameters are then used for deriving catchment rainfall from runoff
757 observations.

758 The cumulative rainfall curves of the different rainfall realisations (ground observation
759 (PObs), INCA (PInca) and inverse rainfall from the different simulation experiments (PInv))
760 are very similar, suggesting, that the inverse model is capable of representing the long-term
761 quantitative rainfall conditions of the catchment. About 60 % (Schliefau) and 45% (Krems) of
762 rainfall is lost to the atmosphere due to actual evapotranspiration (ETa). If the model would
763 not capture ETa adequately, the cumulative rainfall curves would not follow the observations
764 so closely. This is also the case for independent validation periods and years, which were not
765 used in the calibration.

766 The correlation between PInv and PObs, although rather low, is higher or of the same
767 magnitude compared to the correlation between PObs and PInca, suggesting that the inverse
768 model also reflects the timing of rainfall in equal quality of INCA. This is especially the case
769 for the aggregated daily rainfall values. The correlation between PInv and PObs is mostly
770 stable between calibration, validation and in the single years, independent of the simulation
771 experiment. However, again for the simulation experiment with only a single summer for
772 parameter calibration, a larger spread in the correlation for the single years is visible. An
773 increase in observed runoff (Exp4) does not show negative effects on the inverse rainfall
774 measured by the correlation coefficient. A larger bias between observed and modelled rainfall
775 is however visible in Exp4. Generally, the simulation experiment with the highest
776 performance in the runoff simulation also shows the highest correlation values in the rainfall
777 simulations.

778 To test, if the inverse rainfall is conditioned by observed rainfall used as calibration input,
779 additional model calibration is conducted using independent INCA data as driving rainfall
780 input for the forward model calibration. The simulation of inverse rainfall on the basis of this
781 model parameters set show similar results as before, suggesting, that the inverse rainfall is not
782 conditioned to the rainfall input used for model calibration. This result is interesting, since it
783 shows, that the inverse model provides reasonable results in the case where the forward model

784 is calibrated with rainfall data that are independent from the observed rainfall in the proximity
785 of the catchment. Generally, the results do not differ substantially between the two test
786 catchments.

787 Since the inverse model is formulated in a state-space approach additional simulations are
788 performed with differing cold states at the beginning of the simulations. Here the results
789 show, that the inferred rainfall values converge to identical values after an adequate spin-up
790 time.

791 Like with most environmental models, a calibration of the forward model is necessary. It is
792 clear that the application of the inverse model is therefore not possible, if the catchment is
793 completely ungauged. However, this issue is comparable to the application of conventional
794 rainfall-runoff models in gauged and ungauged catchments. As long as a rainfall-runoff model
795 shows reasonable results for the calibration and validation period, the model can be used for
796 different practical applications, e.g. environmental change impact studies, design flood
797 estimations or flood-forecasting. This is also conceivable for the inverse model, since
798 additional information on the catchment rainfall is made available for potential practical
799 applications mentioned above. This additional information is not solely limited to the
800 simulated hourly data, but also includes the aggregated daily rainfall rates, which show a
801 significant higher correlation to the observed values.

802 It can be concluded that the application of the inverse model is a feasible approach to gain
803 additional information on the mean areal rainfall values. The mean areal rainfall values may
804 be used to enhance interpolated rainfall fields, e.g. for the estimation of rainfall correction
805 factors or the parameterisation of elevation dependency. With the inverse model, it is not
806 possible to calculate solid rainfall. In rainless periods, where it is clear, that snow melt is
807 dominating runoff (e.g. in spring), the inverse model can however be used to quantify the
808 snow melt contribution.

809 Areal rainfall estimates leading to extreme flood events are afflicted with major uncertainties.
810 This is underlined by the results where the largest deviations between observed and modelled
811 rainfall is found during flood events. Here the inverse modelling approach can be used as an
812 additional information source concerning the rainfall conditions during extreme events.

813 The inverse model was applied to two catchments. The application and analysis of the
814 proposed method to a wider range of catchments with differing hydrological characteristics is
815 therefore an important task in the near future. Further investigations should include water

limited catchments with an aridity index far lower than 1, where the influences of high evapotranspiration on the inferred rainfall must be investigated.

In the presented work several different model parameter sets were used as a basis to calculate inverse rainfall. In further works the influences and uncertainties in the inverse rainfall, which arise from different model parameters should be analysed systematically. Additionally, a comparison of inverse rainfall estimates from a different model structure for the two catchments with our results would be of interest, in order to check the links between the performance of the forward model and the results obtained by the inversion method.

Appendix

The forward model is formulated as follows, considering parameters and variables in Table 2 and Table 3:

$$BWI_t = \max(\min(INTMAX, BWI_{t-1} + 0.5 * R_t - ETI_t), 0) = \max(\min(INTMAX, BWI_{t-1} + 0.5 * R_t - f(ETp_t, INTMAX)), 0) \quad \text{(A1)}$$

$$R\_Soil_t = 0.5 * R_t + \max(BWI_{t-1} + 0.5 * R_t - ETI_t - INTMAX, 0) \quad \text{(A2)}$$

$$BW0_t = BW0_{t-1} + R\_Soil_t - ETG_t - Q1_t - Q2_t = BW0_{t-1} + R\_Soil_t - \min\left(\frac{BW0_{t-1}}{FKFAK * M}, 1\right) * (ETp_t - ETI_t) * ETVEGCOR - R\_Soil_t * \left(\frac{BW0_{t-1}}{M}\right)^{BETA} - f(PEX2) * BW0_{t-1} \quad \text{(A3)}$$

$$BW2_t = BW2_{t-1} + Q2_t - QAB2_t - QVS2_t = BW2_{t-1} + f(PEX2) * BW0_{t-1} - \alpha_2 * \max(BW2_{t-1} - H2, 0) - \beta_2 * BW2_{t-1} \quad \text{(A4)}$$

$$BW3_t = BW3_{t-1} + QVS2_t - QAB3_t = BW3_{t-1} + \beta_2 * BW2_{t-1} - \alpha_3 * BW3_{t-1} \quad \text{(A5)}$$

$$BW4_t = BW4_{t-1} + Q1_t + QAB2_t + QAB3_t - QSIM_t = BW4_{t-1} + R\_Soil_t * \left(\frac{BW0_{t-1}}{M}\right)^{BETA} + \alpha_2 * \max(BW2_{t-1} - H2, 0) + \alpha_3 * BW3_{t-1} - \alpha_4 * BW4_{t-1} \quad \text{(A6)}$$

with

$$\alpha_i = \frac{\Delta t}{TAB_i} \quad \text{and} \quad \text{(A7)}$$

$$\beta_i = \frac{\Delta t}{TVS_i} \quad \text{(A8)}$$

TAB_i / TVS_i = recession coefficients. Δt = modelling time step in units of hours. α and β vary with modelling time step and represent smoothing functions of the linear reservoirs

Eq. A1 to A8 are simplified representations of the model algorithm. Min/max operators, which, by introducing discontinuities, can lead to non-invertibility. Eq. A4 and A6 do not include a threshold function in the actual model code. The differential equations of the linear reservoirs are solved analytically. An internal time step discretization is included in the code, to guarantee, that the transition between system states above and below the threshold value is

844 solved exactly. A3, representing the soil layer, does include a min() operator for estimating
845 the ratio between actual and potential evapotranspiration as a function of soil water content.
846 This is however not a limiting factor for the inversion, since this factor is a function of the
847 preceding soil state $BW0_{t-1}$, which is known. Only 50% of rainfall is used as input into the
848 interception storage *BWI*. By assuming that the other 50% are always throughfall, eq. A1 and
849 A2 also does not limit the inversion, since a continuous signal through the whole model
850 cascade is guaranteed. The recession coefficient representing percolation processes in the soil
851 layer exhibits a nonlinear characteristic and is calculated as a function of actual soil water
852 content and a as a function of the form parameter PEX2 [-]. This model concept reflects the
853 fact, that higher soil moisture levels lead to higher soil permeability values. These induce
854 higher percolation rates which are reflected by lower recession coefficients.
855

References

Ahrens, B., Jasper, K., and Gurtz, J.: On ALADIN rainfall modeling and validation in an Alpine watershed, Ann. Geophys., 21, 627–637, doi:10.5194/angeo-21-627-2003, 2003.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M.: Crop evapotranspiration: guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. Rome, Italy, 1998.

Bergström, S.: The HBV model, in: Computer Models of Watershed Hydrology, edited by: Singh, V. P., Water Resources Publications, Highland Ranch, CO, USA, 443–476, 1995.

Bica, B., Herrnegger, M., Kann, A., and Nachtnebel, H. P.: HYDROCAST – Enhanced estimation of areal rainfall by combining a meteorological nowcasting system with a hydrological model, Final Report, Austrian Academy of Science, Vienna, doi:10.1553/hydrocast2011, 2011.

BMLFUW: Hydrological Atlas of Austria, 3rd Edn., Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien, ISBN: 3-85437-250-7, 2007.

BMLFUW: Hydrographical yearbook of Austria, Abteilung VII 3 - Wasserhaushalt im Bundesministerium für Land und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien, 2009.

Brent, R.P.: Algorithms for Minimization without Derivatives, Prentice-Hall, Englewood Cliffs, NJ, 1973.

de Jong, C., List, F., and Ergenzinger, C.: Experimental hydrological analyses in the Dischma based on daily and seasonal evaporation, Nord. Hydrol. 33, 1–14, 2002.

Di Baldassarre G., and Montanari, A.: Uncertainty in river discharge observations: a quantitative analysis, Hydrol. Earth Syst. Sci., 13, 913-921, doi:10.5194/hess-13-913-2009, 2009.

Duan, Q., Sorooshian, S., and Gupta, V.K.: Effective and Efficient Global Optimization for Conceptual Rainfall-runoff Models, Water Resour. Res., 28, 1015-1031, 1992.

Eder, G., Fuchs, M., Nachtnebel, H.P., and Loibl, W.: Semi-distributed modelling of the monthly water balance in an alpine catchment, Hydrol. Process. 19, 2339–2360, 2005.

883 Elias, V., Tesar, M., and Buchtele, J.: Occult precipitation: sampling, chemical analysis and
884 process modeling in the Sumava Mts. (Czech Republic) and in the Taunus Mts. (Germany), J.
885 Hydrol., 166, 409–420, 1995.

886 Fekete, B.M., Vorosmarty, C.J., Roads, J.O., and Willmot, C.J.: Uncertainties in precipitation
887 and their impacts on runoff estimates, J. Clim. 17, 294– 304, 2004.

888 Goodison, B.E., Louie, P.Y.T., and Yang, D.: WMO solid precipitation measurement
889 intercomparison, Instruments and Observing Methods Rep. 67 (WMO/TD 872), World
890 Meteorological Organization, Geneva, Switzerland, 318 pp, 1998.

891 Groetsch, C.: Inverse Problems in Mathematical Sciences, Vieweg Mathematics for Scientists
892 and Engineers, Wiesbaden, 1993.

893 Haiden, T., and Pistotnik, G.: Intensity-dependent parameterization of elevation effects in
894 precipitation analysis, Adv. Geosci., 20, 33–38, doi:10.5194/adgeo-20-33-2009, 2009.

895 Haiden, T., Kann, A., Wittman, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated
896 Nowcasting through Comprehensive Analysis (INCA) system and its validation over the
897 Eastern Alpine region, Weather Forecast., 26, 166–183, doi:10.1175/2010WAF2222451.1,
898 2011.

899 Hargreaves, G.H., and Samani, Z.A.: Estimating potential evapotranspiration, J. Irr. Drain.
900 Div-ASCE, 108, 225–230, 1982.

901 Herrnegger, M.: Zeitlich hochaufgelöste inverse Modellierung von Gebietsniederschlägen aus
902 Abflussmessungen, PhD thesis, Institute of Water Management, Hydrology and Hydraulic
903 Engineering, University of Natural Resources and Life Sciences, Vienna, Austria, 2013.

904 Herrnegger, M., Nachtnebel, H.P., and Haiden, T.: Evapotranspiration in high alpine
905 catchments - an important part of the water balance!, Hydrol. Res. 43, 460-475, 2012.

906 Hino, M., and Hasabe, M.: Analysis of hydrologic characteristics from runoff data – a
907 hydrologic inverse problem. J. Hydrol., 49, 287-313, 1981.

908 Jacobs, A.F.G., Heusinkveld, B.G., and Wichink Kruit, R.J.: Contribution of dew to the water
909 budget of a grassland area in the Netherlands, Water Resour. Res., 42, W03415,
910 doi:10.1029/2005WR004055, 2006.

911 Jasper, K. and Kaufmann, P.: Coupled runoff simulations as validation tools for atmospheric
912 models at the regional scale. Q. J. Roy. Meteorol. Soc., 129, 673-692, 2007.

913 Jasper, K., Gurtz, J., and Lang, H.: Advanced flood forecasting in Alpine watersheds by
914 coupling meteorological observations and forecasts with a distributed hydrological model, J.
915 Hydrol., 267, 40-52, 2002.

916 Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization,
917 rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45, W02429,
918 doi:10.1029/2008WR006912, 2009.

919 Klemm, O., and Wrzesinski, T.: Fog deposition fluxes of water and ions to a mountainous site
920 in Central Europe, Tellus 59, 705-714, 2007.

921 Kling, H.: Spatio-temporal modelling of the water balance of Austria. Dissertation, University
922 of Natural Resources and Applied Life Sciences, 234 pp., available at:
923 http://iwhw.boku.ac.at/dissertationen/kling.pdf (last access: 7 October 2014), 2006.

924 Kling, H., and Nachtnebel, H.P.: A method for the regional estimation of runoff separation
925 parameters for hydrological modelling, J. Hydrol., 364, 163–174, 2009.

926 Kling, H., Stanzel, P., Fuchs, M., and Nachtnebel, H.P.: Performance of the COSERO
927 precipitation-runoff model under non-stationary conditions in basins with different climates,
928 Hydrolog. Sci. J., doi: 10.1080/02626667.2014.959956, 2015.

929 Krajewski, W.F., and Smith, J.A.: Radar hydrology: rainfall estimation, Adv. Water Resour.,
930 25, 1387-13, 2002.

931 Krajewski, W.F., Villarini, G., and Smith, J.A.: RADAR-rainfall uncertainties, B. Am.
932 Meteorol. Soc., 91, 87–94. doi:10.1175/2009BAMS2747.1, 2010

933 Krier, R., Matgen, P., Goergen, K., Pfister, L., Hoffmann, L., Kirchner, J. W., Uhlenbrook, S.,
934 and Savenije, H.H.G.: Inferring catchment precipitation by doing hydrology backward: A test
935 in 24 small and mesoscale catchments in Luxembourg, Water Resour. Res., 48, W10525,
936 doi:10.1029/2011WR010657, 2012.

937 Kuczera, G., Kavetski, D., Franks, S., and Thyer, M.: Towards a Bayesian total error analysis
938 of conceptual rainfall-runoff models: Characterising model error using storm-dependent
939 parameters, J. Hydrol., 331, 161–177, 2006.

940 Kunstmann, H., and Stadler, C.: High resolution distributed atmospheric-hydrological
941 modeling for Alpine catchments, J. Hydrol., 314, 105-124, 2005.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. Hydrol. Earth Syst. Sci., 16, 3863-3887, 2012.

McLaughlin, D.: An integrated approach to hydrologic data assimilation: interpolation, smoothing and filtering. Advances in Water Resources, 25, 1275-1286, 2002.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., and Clark, M.: Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, Hydrol. Process., 24, 1270–1284, doi:10.1002/Hyp.7587, 2010.

Nachtnebel, H. P., Herrnegger, M., Kahl, B., and Hepp, G.: Meteorologisch-hydrologisches Warnsystem Steyr: Endbericht und Technische Dokumentation - Teil 3 - Hydrologische Abflussmodellierung, Amt der OÖ Landesregierung - Abteilung Wasserwirtschaft, Schutzwasserwirtschaft und Hydrographie, 197, 2010a.

Nachtnebel, H.P., Senoner, T., Kahl, B., Apperl, B., and Waldhör, B.: Hochwasserprognosesystem Ybbs - Hydrologische Abflussmodellierung, Amt der NÖ Landesregierung, St. Pölten, 176, 2010b.

Nachtnebel, H.P., Haberl, U., Stanzel, Ph., Kahl, B., Holzmann, H., and Pfaffenwimmer, Th.: Hydrologische Abflussmodellierung - Teil 3, in: Amt der Salzburger Landesregierung: HydrisII Hydrologisches Informationssystem zur Hochwasservorhersage im Land Salzburg, Amt der Salzburger Landesregierung, 341, 2009a.

Nachtnebel, H. P., Senoner, T., Stanzel, P., Kahl, B., Hernegger, M., Haberl, U. and Pfaffenwimmer, T.: Inflow prediction system for the Hydropower Plant Gabčíkovo, Part 3 - Hydrologic Modelling, Slovenské elektrárne, a.s. Bratislava, 139, 2009b.

Nachtnebel, H.P., Baumung, S., and Lettl, W.: Abflussprognosemodell für das Einzugsgebiet der Enns und Steyr, report, Institute of Water Management, Hydology and Hydraulic Engineering, University of Natural Resources and Applied Life Sciences, Vienna, Austria, 1993.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models, Part I: A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Pappenberger, F., Matgen, P., Beven, K.J., Henry J.B., Pfister, L., and de Fraipont, P.: Influence of uncertain boundary conditions and model structure on flood inundation predictions, Adv. Water Resour., 29, 1430–1449, 2006.

Pelletier, M.P.: Uncertainties in the determination of river discharge: a literature review, Can. J. Civ. Eng., 15, 834–850, 1987.

Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, J. Hydrol., 242, 275–301, 2001.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P.: Numerical Recipes in FORTRAN, The Art of Scientific Computing, Cambridge Univ. Press, New York, 965 pp., 1992.

Seibert, J.: Reliability of model predictions outside calibration conditions, Nord. Hydrol., 34, 477–492, 2003.

Seibert, J., and Morén, A.-S.: Reducing systematic errors in rainfall measurements using a new type of gauge, Agric. Forest. Meteorol., 98–99, 341–348, 1999.

Sevruk, B.: Methodische Untersuchungen des systematischen Messfehlers der Hellmann-Regenmesser im Sommerhalbjahr in der Schweiz, Dissertation, Eidgenöss. Techn. Hochsch. Zürich, Zürich, Switzerland, 1981.

Sevruk, B.: Correction of precipitation measurements. Proc. Workshop on the Correction of Precipitation Measurements, in: Zürcher Geographische Schriften, ETH Zürich, p. 289, 1986.

Sevruk, B., and Nespor, V.: Empirical and theoretical assessment of the wind induced error of rain measurement, Water Sci. Technol., 37, 171–178, 1998.

Simoni, S., Padoan, S., Nadeau, D.F., Diebold, M., Porporato, A., Barrenetxea, G., Ingelrest, F., Vetterli, and M., Parlange, M.B.: Hydrologic response of an alpine watershed: application of a meteorological wireless sensor network to understand streamflow generation, Water Resour. Res., 47, W10524, doi:10.1029/2011WR010730, 2011.

Stanzel, Ph., Kahl, B., Haberl, U., Herrnegger, M., and Nachtnebel, H. P.: Continuous hydrological modeling in the context of real time flood forecasting in alpine Danube tributary catchments, IOP Conference Series, 4, 012005, doi:10.1088/1755-1307/4/1/01200, 2008.

1001 Sugawara, M.: On the weights of precipitation stations. in: O'Kane, J.P. (Ed.), Advances in
1002 Theoretical Hydrology, edited by O´Kane, J.P., Elsevier Science Publishers, Amsterdam, 59–
1003 74, 1992.

1004 Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Society
1005 for Industrial and Applied Mathematics, Philadelphia, 352 pp., 2005.

1006 Valéry, A., Andréassian, V., and Perrin, C.: Inverting the hydrological cycle: when
1007 streamflow measurements help assess altitudinal precipitation gradients in mountain areas, in:
1008 New Approaches to Hydrological Prediction in Data–sparse Regions, IAHS Publ., 333, 281-
1009 286, 2009.

1010 Valéry, A., Andréassian, V., and Perrin, C.: Regionalisation of rainfall and air temperature
1011 over high-altitude catchments – learning from outliers, Hydrol. Sci. J., 55, 928–940, 2010.

1012 van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of
1013 unsaturated soils, Soil Sci. Soc. Am J., 44, 892-898, 1980.

1014 Vrugt, J.A., ter Braak, C.J.F., Clark, M.P., Hyman, J.M., and Robinson B.A.: Treatment of
1015 input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain
1016 Monte Carlo simulation, Water Resour. Res., 44, W00B09, doi:10.1029/2007WR006720,
1017 2008.

1018 Wood, S.J., Jones, D.A., and Moore, R.J.: Accuracy of rainfall measurement for scales of
1019 hydrological interest, Hydrol. Earth Syst. Sci., 4, 531–543, doi:10.5194/hess-4-531-2000.

1020

1021

1022 **Tables**

1023 Table 1: Magnitude of different systematic errors in precipitation measurements (Sevruk,
1024 1981, 1986; Goodison et al, 1998; Elias et al., 1993; Jacobs et al., 2006; Klemm and
1025 Wrzesinsky, 2007).

Systematic error	Magnitude
Wind-induced errors	2 - 10 % (liquid precipitation) 10 - >50 % (snow)
Wetting losses	2 - 10 %
Evaporation losses	0 - 4 %
Splash-out and splash-in	1 - 2 %
Fog and dew	4 - 10 %

1026

1027 Table 2: Model parameters θ_i. Parameters in *italics* are calibrated.

Parameter	Units	Range	Description
INTMAX	mm	0.5 - 2.5	Interception storage capacity
M	mm	80 - 250	Soil storage capacity
FKFAK	-	0.5 - 1	Critical soil moisture for actual evapotranspiration
ETVEGCOR	-	0.4 - 1.1	Vegetation correction factor for actual evapotranspiration from soil
BETA	-	0.1 - 10	Exponent for computing fast runoff generation
KBF	h	4000 - 12000	Recession coefficient for percolation from soil module
PEX2	-	5 - 25	Parameter for non-linear percolation
TAB2	h	50 - 500	Recession coefficient for interflow
TVS2	h	50 - 500	Recession coefficient for percolation from interflow reservoir
H2	mm	0 - 25	Outlet height for interflow
TAB3	h	1000 - 5000	Recession coefficient for base flow
TAB4	h	0.05 - 10	Recession coefficient for routing

1028

1029

1030 Table 3: Model fluxes and system states S_i. Fluxes represent sums over the time step.

Variable	Units	Type	Description
R	mm	Input	Rainfall
ETp	mm	Input	Potential evapotranspiration
ETI	mm	Output	Actual evapotranspiration from interception module
ETG	mm	Output	Actual evapotranspiration from soil module
BWI	mm	State	Water stored in interception module
BW0	mm	State	Water stored in soil module
BW2	mm	State	Water stored in interflow reservoir
BW3	mm	State	Water stored in base flow reservoir
BW4	mm	State	Water stored in routing reservoir
R_Soil	mm	Internal flux	Input into soil module
Q1	mm	Internal flux	Fast runoff from soil module
Q2	mm	Internal flux	Percolation from soil module
QAB2	mm	Internal flux	Interflow
QVS2	mm	Internal flux	Percolation from interflow reservoir
QAB3	mm	Internal flux	Base flow
QSIM	mm	Output	Total runoff

1031

1032 Table 4: Characteristics of the study catchments (BMLFUW, 2007; BMLFUW, 2009).

	Schliefau	Krems
Basin area [km²]	17.9	38.4
Mean elevation [m]	608	598
Elevation range [m]	390 - 818	413 - 1511
Mean annual precipitation [mm]	1390	1345
Mean annual runoff [m³/s]	0.38	1.12

1033

1034

Table 5: Overview of the model calibration and simulations experiments with observed input data. POobs and PInca refer to the rainfall from the station observations and the INCA system.

	Jun. to Sept. in year				Driving input (Forward / inverse model)	Purpose
	2006	2007	2008	2009		
Exp1	calib.	valid.	valid.	valid.	POobs / Q	Influence of different calibration periods on simulations
Exp2	calib.	calib.	valid.	valid.	POobs / Q	
Exp3	calib.	calib.	calib.	valid.	POobs / Q	
Exp4	calib.	calib.	calib.	valid.	POobs / Q+10%	Influence of different runoff Q on simulations
Exp5	calib.	calib.	calib.	valid.	PInca / Q	Influence of different rainfall input on simulations
Exp6	Parameters from Exp3, but different initial conditions				POobs / Q	Influence of cold states on simulations

Table 6: Correlation (CORR), mean bias and mean squared error (MSE) for different temporal aggregation lengths between observed and inferred precipitation of the virtual experiments, in which different magnitudes of noise was added to the input runoff data. These are indicated with the "Noise scaling factor".

Noise scaling factor	CORR [-]			Mean BIAS [mm]			MSE [mm²]		
	1h-sums	6h-sums	24h-sums	1h-sums	6h-sums	24h-sums	1h-sums	6h-sums	24h-sums
0%	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
1%	0.994	0.999	1.000	-0.001	-0.007	-0.028	0.011	0.016	0.015
2%	0.982	0.998	1.000	-0.003	-0.015	-0.060	0.034	0.051	0.043
5%	0.921	0.991	0.999	-0.007	-0.040	-0.160	0.154	0.300	0.230
10%	0.819	0.977	0.998	-0.013	-0.079	-0.316	0.408	0.770	0.556

Table 7: Model performance for the different simulation experiments and the two catchments of the forward model, expressed by Nash-Sutcliffe-Efficiency (NSE) and the mean bias between simulated and observed runoff in percent of observed runoff for different periods. Only the months June to September are evaluated.

		NSE [-]			mean Bias [%]		
		Calib.	Valid.	2006-2009	Calib.	Valid.	2006-2009
Schliefau	Exp1	0.872	0.814	0.822	4.8	8.7	7.8
	Exp2	0.858	0.819	0.832	11.4	-0.8	3.9
	Exp3	0.812	0.837	0.828	1.5	0.1	0.9
	Exp4	0.814	0.840	0.830	-4.4	-8.3	-5.9
	Exp5	0.738	0.715	0.728	2.1	-4.9	-0.6
Krems	Exp1	0.879	0.740	0.763	-9.4	1.2	-1.4
	Exp2	0.849	0.851	0.851	-0.3	-8.6	-4.8
	Exp3	0.842	0.855	0.851	-3.2	-8.0	-4.8
	Exp4	0.845	0.859	0.854	-6.1	-11.5	-7.9
	Exp5	0.748	0.815	0.787	3.7	-2.8	1.5

Table 8: Correlation between different rainfall realisations, evaluated for different periods and for 1h- and 24-h-sums. (PObs: Ground observation, PInv: Inverse rainfall from Exp1 to Exp5, PInca: INCA rainfall).

		CORR: 1h-sums					CORR: 24h-sums				
		PObs - PInv			PInca-PInv	PObs - PInca	PObs - PInv			PInca-PInv	PObs - PInca
		Calib.	Valid.	2006-2009	2006-2009	2006-2009	Calib.	Valid.	2006-2009	2006-2009	2006-2009
Schliefau	Exp1	0.706	0.460	0.504	0.251		0.935	0.857	0.871	0.802	
	Exp2	0.572	0.540	0.549	0.290		0.939	0.895	0.914	0.840	
	Exp3	0.515	0.567	0.534	0.284	0.463	0.913	0.929	0.918	0.845	0.928
	Exp4	0.515	0.558	0.530	0.283		0.910	0.928	0.917	0.843	
	Exp5	0.514	0.545	0.524	0.276		0.916	0.927	0.920	0.842	
Krems	Exp1	0.622	0.430	0.478	0.394		0.880	0.871	0.871	0.847	
	Exp2	0.437	0.602	0.517	0.445		0.907	0.910	0.909	0.889	
	Exp3	0.493	0.581	0.517	0.445	0.469	0.896	0.936	0.909	0.889	0.931
	Exp4	0.494	0.577	0.517	0.445		0.896	0.936	0.909	0.892	
	Exp5	0.473	0.593	0.503	0.445		0.884	0.936	0.901	0.888	

Table 9: Mean Bias in mm between different rainfall realisations, evaluated for different periods and 1h- and 24h-sums.

		Mean Bias: 1h-sums [mmh^{-1}]					Mean Bias: 24h-sums [mmd^{-1}]				
		PInv - PObs			PInv - PInca	PInca - PObs	PInv - PObs			PInv - PInca	PInca - PObs
		Calib.	Valid.	2006-2009	2006-2009	2006-2009	Calib.	Valid.	2006-2009	2006-2009	2006-2009
Schliefau	Exp1	0.001	0.007	0.006	0.005		0.019	0.179	0.139	0.118	
	Exp2	-0.008	0.014	0.003	0.002		-0.204	0.339	0.067	0.046	
	Exp3	0.003	0.027	0.009	0.008	0.001	0.075	0.639	0.216	0.195	0.021
	Exp4	0.009	0.041	0.017	0.016		0.225	0.986	0.415	0.394	
	Exp5	0.007	0.034	0.014	0.013		0.169	0.817	0.331	0.310	
Krems	Exp1	0.029	0.006	0.012	-0.008		0.686	0.148	0.283	-0.191	
	Exp2	0.013	0.020	0.017	-0.003		0.324	0.485	0.404	-0.069	
	Exp3	0.015	0.022	0.017	-0.003	0.020	0.362	0.531	0.404	-0.069	0.473
	Exp4	0.019	0.033	0.022	0.003		0.450	0.785	0.534	0.061	
	Exp5	0.020	0.022	0.021	0.001		0.478	0.536	0.493	0.019	

Figures

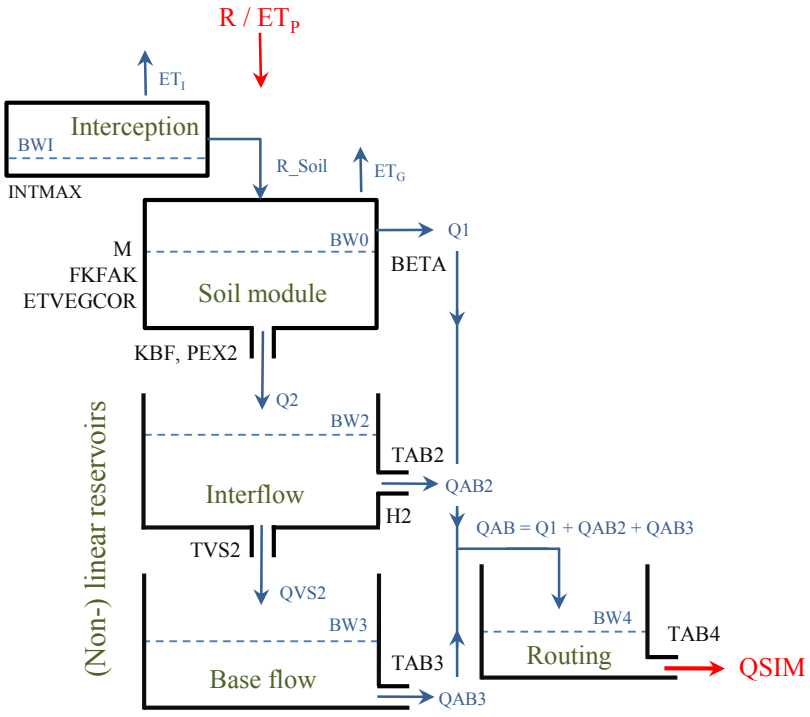


Figure 1: Structure, parameters and states of the forward model.

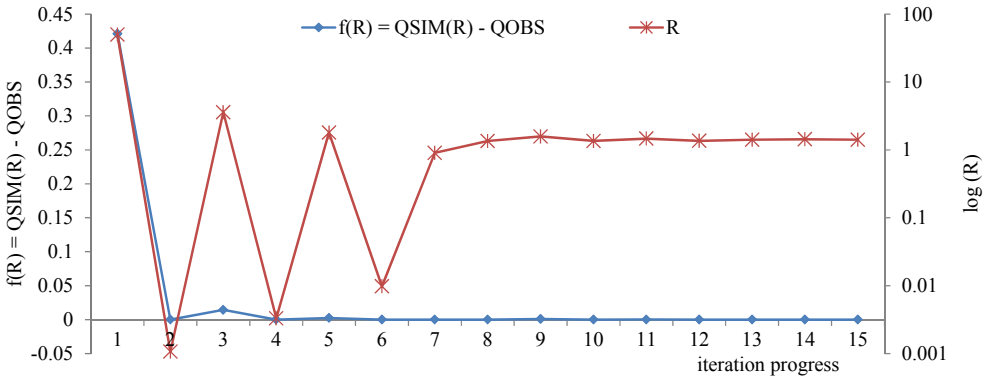
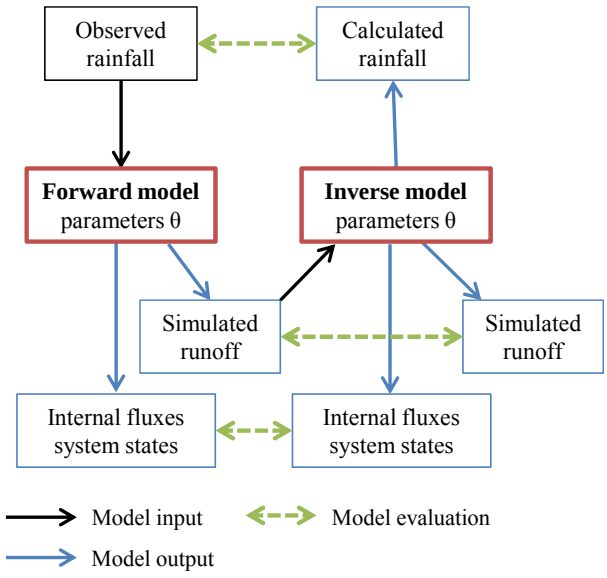


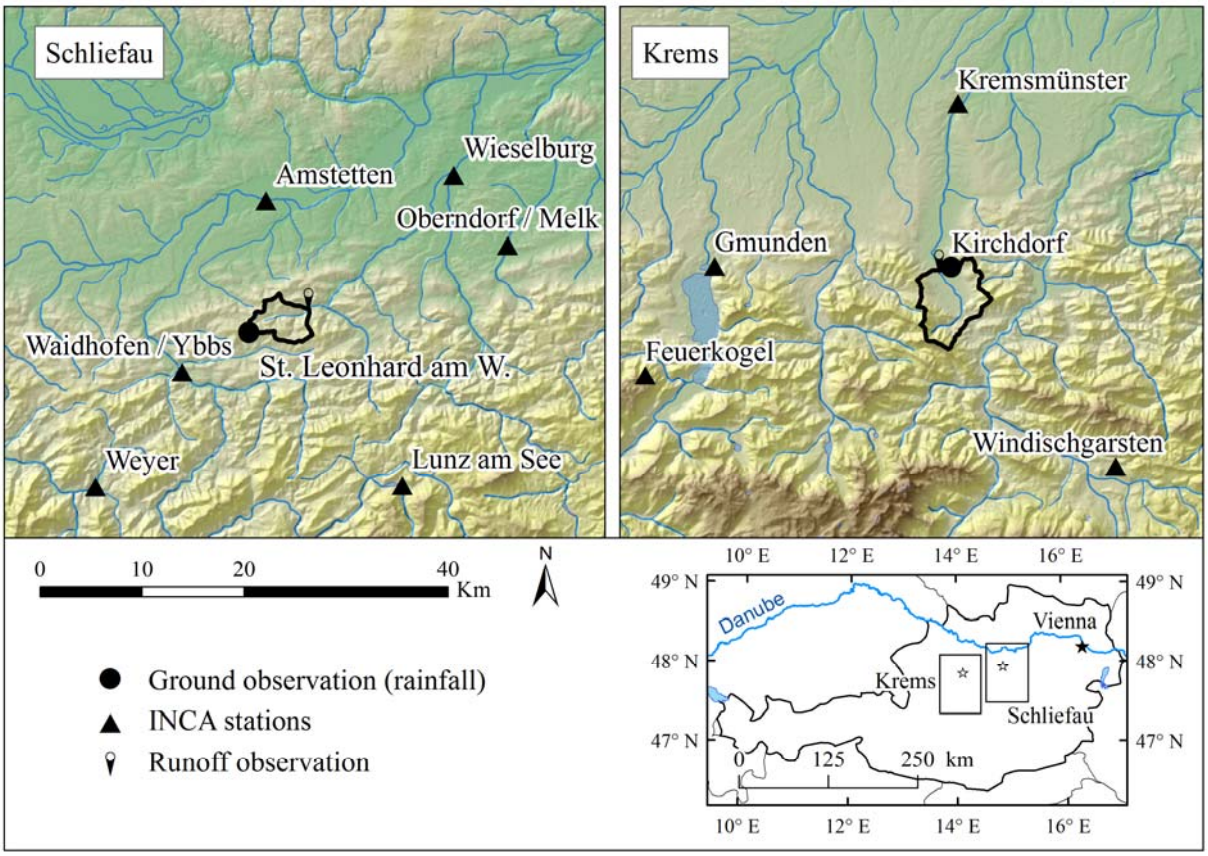
Figure 2: Illustration of the iteration progress for one model time step. Note that the right y-axis showing the inverse rainfall values (R) is in a logarithmic scale (units in mm/h).



1067

1068 Figure 3: Setup of the virtual experiments and evaluation of the inverse model. All variables
1069 are calculated for every Monte Carlo run, in which parameters θ are varied.

1070



1071

1072 Figure 4: Schliefau and Krems catchment and location of meteorological stations. Note that
1073 ground observation of rainfall is not part of the INCA stations network.

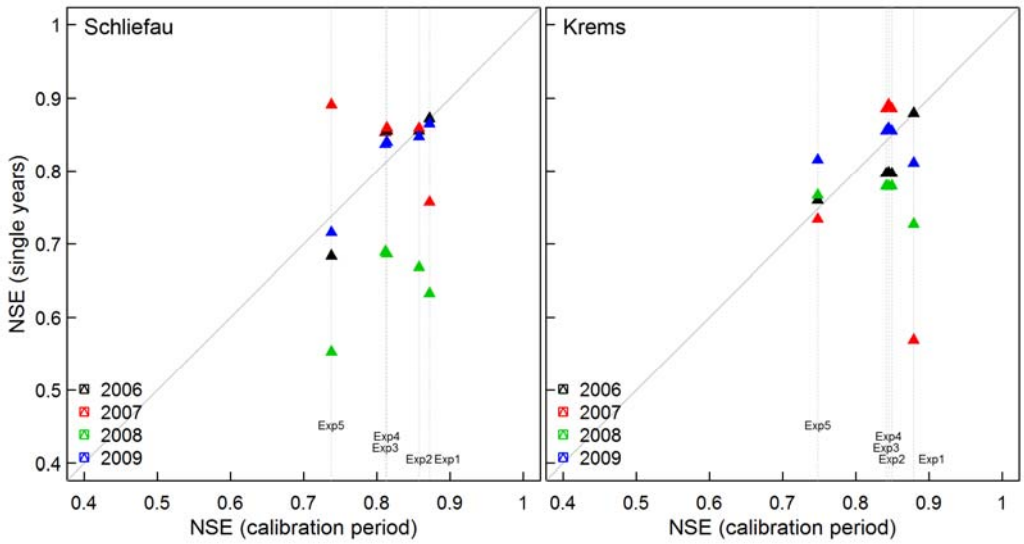


Figure 5: Nash-Sutcliffe-Efficiency (NSE) of the forward model for the calibration periods versus single years for the two study areas.

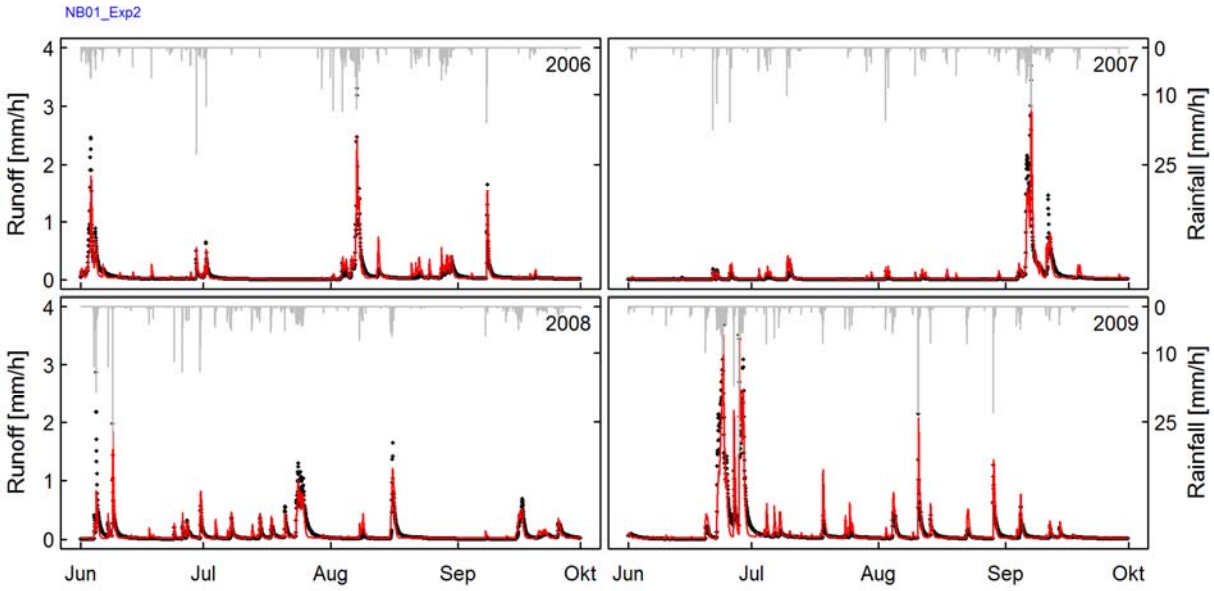


Figure 6: Schliefau catchment: Observed (black points) and simulated (red) runoff of Exp2.

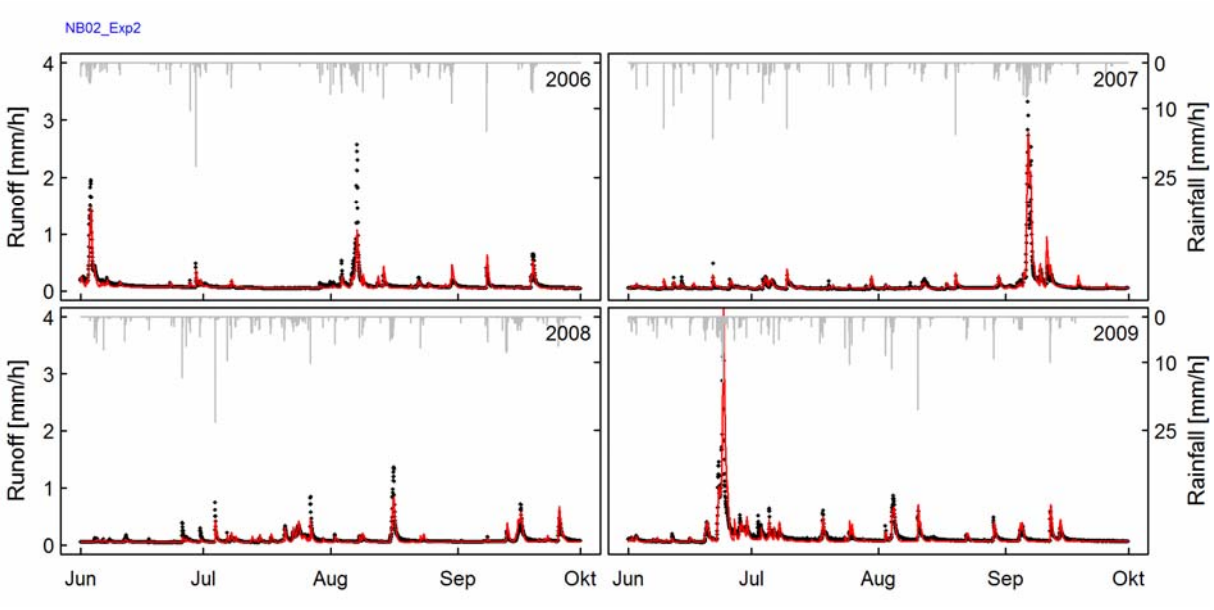
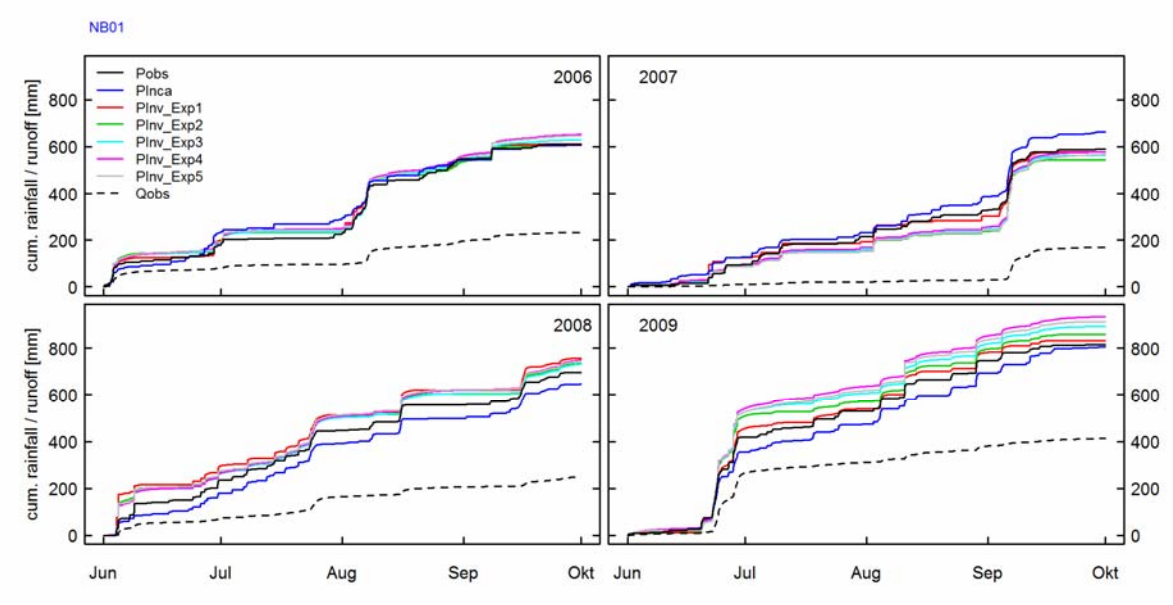


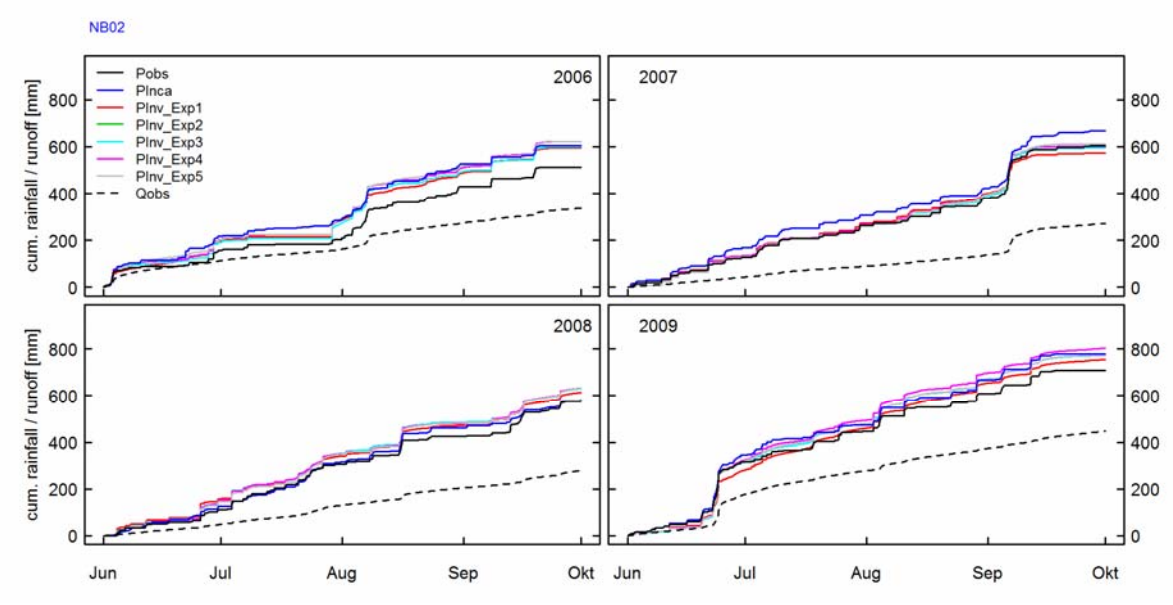
Figure 7: Krems catchment: Observed (black points) and simulated (red) runoff of Exp2.



1087

1088 Figure 8: Schliefau catchment: Cumulative rainfall curves for observed rainfall (PObs), INCA
1089 rainfall (PInca) and the inverse rainfall of Exp1 to Exp5 (PInv). Cumulative sums of observed
1090 runoff are shown as dashed black lines.

1091



1092

1093 Figure 9: Krems catchment: Cumulative rainfall curves for observed rainfall (PObs), INCA
1094 rainfall (PInca) and the inverse rainfall of Exp1 to Exp5. Cumulative sums of observed runoff
1095 are shown as dotted black lines.

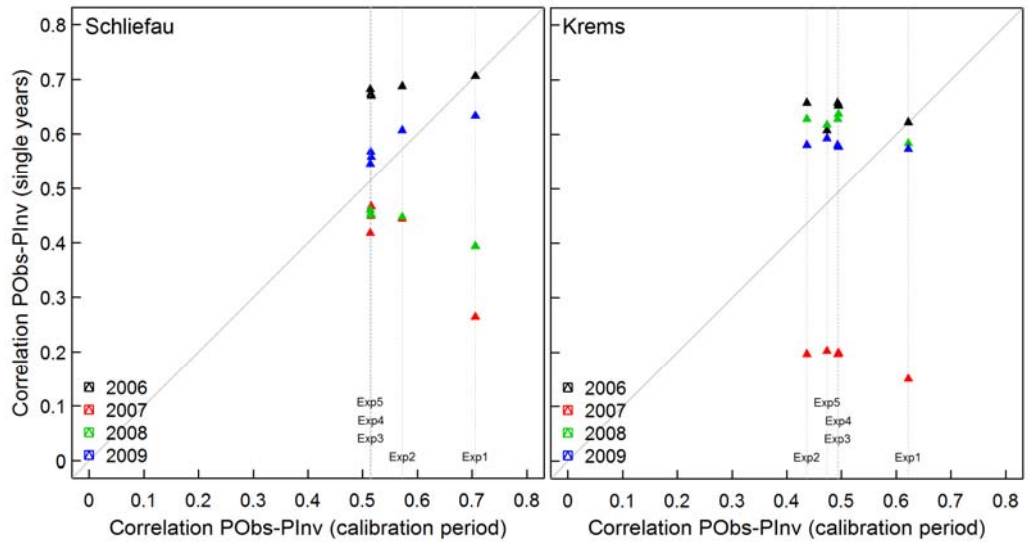


Figure 10: Correlation between POobs-PInv for the calibration periods of the simulation experiments Exp1 to Exp5 versus single years for the two study areas.

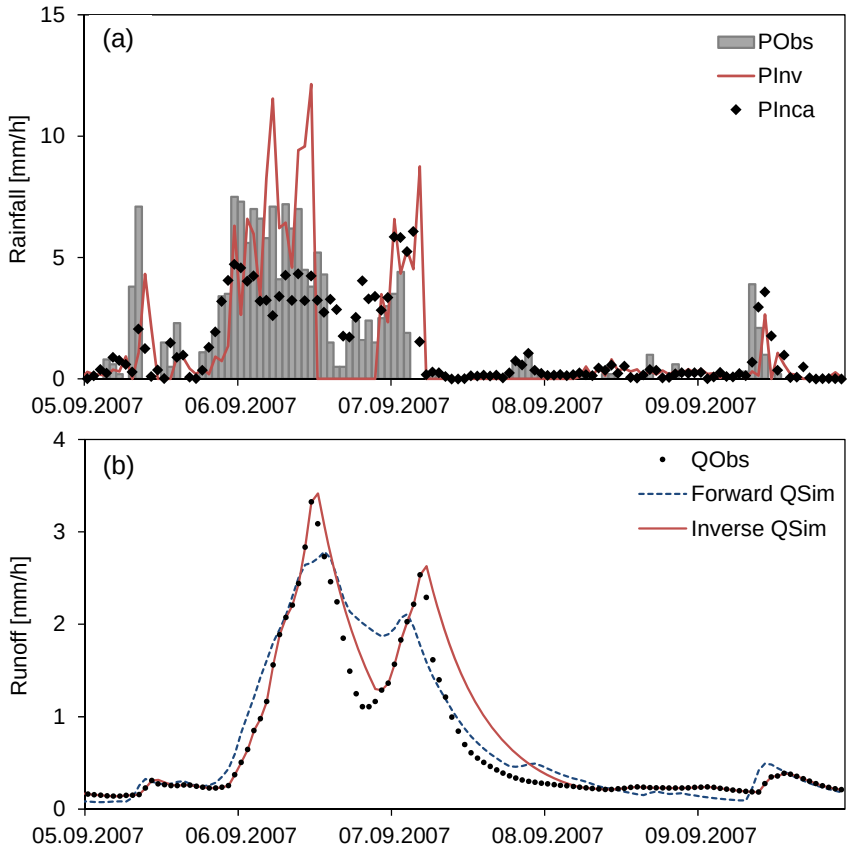
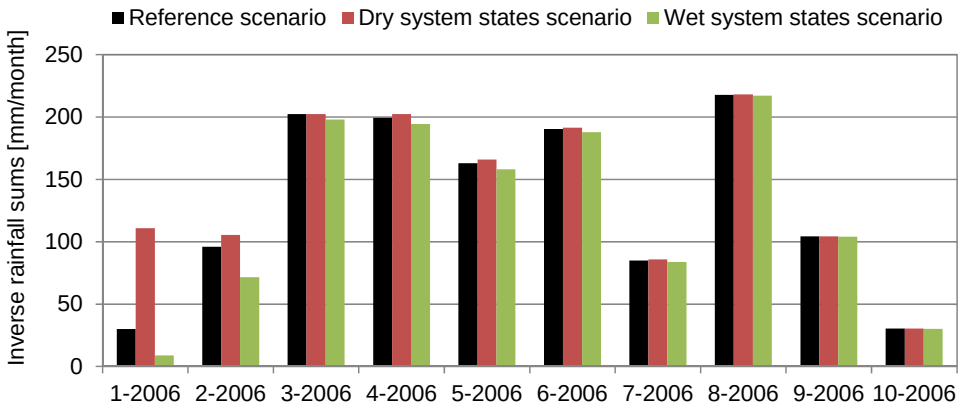


Figure 11: Krems catchment: Temporal development of the different rainfall realisations (a) and runoff (b) for a flood event. Simulations originate from Exp3.



1102

1103 Figure 12: Krems catchment: Monthly sums of inverse rainfall simulated in the scenarios
1104 "reference", "dry" and "wet" from Exp6.