

Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany)

K. Förster^{1,2,3}, G. Meon¹, T. Marke², and U. Strasser²

¹Department of Hydrology, Water Management and Water Protection, Leichtweiss Institute for Hydraulic Engineering and Water Resources, University of Braunschweig, Beethovenstr. 51a, 38106 Braunschweig, Germany

²Institute of Geography, University of Innsbruck, Innrain 52f, 6020 Innsbruck, Austria

³alpS, Grabenweg 68, 6020 Innsbruck, Austria

Correspondence to: K. Förster (kristian.foerster@uibk.ac.at)

Abstract

Detailed physically based snow models using energy balance approaches are spatially and temporally transferable and hence regarded as particularly suited for scenario applications including changing climate or land-use. However, these snow models place high demands on meteorological input data at the model's scale. Besides precipitation and temperature, time series of humidity, wind speed, and radiation have to be provided. In many catchments these time series are rarely available or provided by a few meteorological stations only. This study analyses the effect of improved meteorological input on the results of four snow models with different complexity for the Sieber catchment (44.4 km²), Harz Mountains, Germany. The Weather Research and Forecast model (WRF) is applied to derive spatial and temporal fields of meteorological surface variables at hourly temporal resolution for a regular grid of 1.1 × 1.1 km². All snow models are evaluated at the point and the catchment scale. For catchment scale simulations all snow models were integrated into the hydrological modelling system PANTA RHEI. The model results achieved with a simple temperature-index model using observed precipitation and temperature time series as input are compared to those achieved with WRF input. Due to a mismatch between modelled and observed precipitation, the observed melt runoff as provided by a snow lysimeter and the observed streamflow are better reproduced by application of observed meteorological input data. In total, precipitation is simulated statistically reasonably at the seasonal scale but some single precipitation events are not captured by the WRF dataset. Regarding the model efficiencies achieved for all simulations using WRF data, energy balance approaches generally perform similar compared to the temperature-index approach and partially outperform the latter.

1 Introduction

The water balance of mountainous or sub-arctic catchments is strongly influenced by snow processes. Due to accumulation of snow during the winter season, a considerable fraction of precipitation is stored in the snow pack. Rapid snowmelt, partially superimposed by rain events,

can cause extensive flooding with high potential of damage. Hence, hydrological modelling systems have to incorporate snow accumulation and melt for water resources management in regions that are affected by snow processes. In general, two types of snow models are available:

- Temperature-index model (also known as degree-day method) and extended index models (e.g., Anderson, 1973): these models are “simple and modest” (Rango and Martinec, 1995) because only temperature and precipitation time series are needed. Some extensions also incorporate shortwave radiation input (Hock, 1999; Pellicciotti et al., 2005).
- Energy balance models include a detailed physically based description for the surface energy balance of a snow pack (e.g., Anderson, 1968). The process based equations require further meteorological time series as input (e.g., humidity, radiation, wind speed).

When using energy balance approaches instead of index melt models, many more processes such as sublimation losses can be considered. These can be relevant for the Alpine water balance (Strasser et al., 2008). Even though some parameterizations of the energy balance can also be considered inductive empirical functions (Beven, 2012), energy balance approaches are considered to be more reliable for scenario simulations than simpler index methods due to their physical basis (e.g., Charbonneau et al., 1981; Walter et al., 2005; Pomeroy et al., 2007; Singh et al., 2009; Barry and Gan, 2011; Warscher et al., 2013; Marke et al., 2014b).

There are many physically based snow models available for point scale applications (see, e.g., Etchevers et al., 2004; Rutter et al., 2009). Nevertheless, the temperature-index method is widely accepted in hydrological modelling at the catchment scale (Rango and Martinec, 1995; Seibert, 1999; Beven, 2001). Simple index approaches outperform energy balance approaches in many cases because the latter are more sensitive to the quality of meteorological input data (Zappa et al., 2003). Hence, “inadequate basin-scale hydrologic observations” (Franz et al., 2008) restrict catchment-scale applications of energy balance approaches. However, some studies prove that energy balance approaches provide better results than temperature-index models, even at the catchment scale (see, e.g., Kuchment and Gelfan, 1996; Fuka et al., 2012).

Authors like Bales et al. (2006), Franz (2006), and Jost et al. (2012) suggest finding new strategies to provide suitable data to drive process-based snow models for catchment-scale

applications. For instance, Franz (2006) and El-Sadek et al. (2011) propose (re-)analysis data as input for hydrological modelling. However, this type of data does not provide a sufficient spatial and temporal resolution for typical applications in hydrological modelling for complex topographies of mountain regions. In consideration of topography, land use and soil type, regional models of the atmosphere that are also referred to as limited-area models (LAM) are suitable tools to “add regional detail” to these global scale (re-)analysis data (Giorgi, 2006). This procedure is called dynamical downscaling. Besides hydrological forecasts and climate impact studies, successful applications of LAM in hydrological modelling and cryospheric research using (re-) analysis data were reported by Kunstmann and Stadler (2005), Rögnvaldsson et al. (2007a), Rögnvaldsson et al. (2007b), Bernhardt et al. (2010), Liu et al. (2011), Marke et al. (2011), Maussion et al. (2011), Pavelsky et al. (2011), Mölg et al. (2012), and Marke et al. (2014a). In addition, globally available datasets can also be considered valuable information for predictions in ungauged basins (Sivapalan et al., 2003) and ungauged climates (Merz et al., 2011).

The main questions to be answered in this study are:

- To what extent does downscaled LAM data enhance model performance?
- Does increasing snow model complexity using these data increase model performance?

A LAM of the atmosphere is applied to obtain hourly spatially distributed meteorological time series for snowmelt simulations at both the point and the catchment scale. It is assumed that the meteorological forcing data generated by a LAM, by means of dynamical downscaling, are physically consistent in space and time (Giorgi, 2006) – a basic requirement for application as input for physically based snow models. As reported in the above-cited studies, analysis data are used herein as input for the LAM. In contrast or extension to other studies that address LAM input, four independent snow models including the temperature-index model and three energy balance models are applied in order to investigate different degrees of model complexity on the quality snow model results. Moreover, using three energy balance models representing more complex models is seen as a multiple hypothesis framework (Clark et al., 2011). Since most energy balance approaches have been developed for point-scale applications but hydrological

models requires upscaling of these models, all snow models are tested at the point and at the catchment scale.

2 Data and methods

2.1 Study area

5 All simulations were carried out in the Harz Mountains, a low mountain range in the northern part of Germany (see Fig. 1). The study area covers elevations ranging from 300 to 1100 m a.s.l. The Brocken (1142 m a.s.l.) is the highest peak of both the Harz Mountains and northern Germany. The Harz Mountains delineate the northern boundary of low mountain ranges in Germany merging into the North European Plain with elevations below 200 m a.s.l.

10 Due to altitudinal differences, considerable gradients in meteorological fields reflecting different climates exist. The mean annual temperature and precipitation depth at Clausthal climate station (585 m a.s.l.) are 6.2 °C and 1326 mm a⁻¹, respectively. However, the climate at the Brocken is considerably different (2.9 °C, 1814 mm a⁻¹). All values provided above refer to the period 1961–1990.

15 Two sites were selected for the simulations in this study:

- Torfhaus meteorological station (805 m a.s.l.) for point-scale applications: Besides precipitation, temperature and snowdepth recordings, a snowmelt lysimeter is available to provide melt rates. This snowmelt lysimeter is an unenclosed type with a small rim above the collector, which covers an area of 2 m². A tipping bucket is installed below the collector in order to continuously record melt rates. Based on recordings from 2004 to 2012 the mean annual precipitation depth can be approximated as 1430 mm a⁻¹.
- Sieber catchment upstream of Pionierbrücke gauging station (340 m a.s.l.) for catchment-scale applications: The upstream Sieber catchment covers an area of 44.4 km² with elevations ranging from 340 m to 920 m a.s.l. Soil types of sandy loam and loamy sand over bedrock are prevailing. About 75 % of the catchment is covered by coniferous forest.

25

Norway spruce (*Picea abies*) is the predominant wood species. The remaining parts of the area are covered by deciduous trees, meadows, upland moors, and minor settlements. Due to former mining activities in the Harz Mountains, a system of channels redirects water across the watershed boundary (Upper Harz Water Management System). The mean annual runoff depth, which was calculated for the period 1930–2013 based on long-term observations, is 1097 mm a⁻¹. Even though average monthly precipitation depth is highest in December, the maximum value of mean monthly runoff depth typically occurs in April. This shift between the mean annual courses of precipitation and runoff depth can be related to snow accumulation during the winter and snowmelt in spring.

2.2 Selected winter seasons

As will be described later in Sect. 2.3 in much more detail, computation time still limits the application of LAMs with respect to the number of runs, spatial resolution or length of LAM runs (Rummukainen, 2010). Therefore, apart from large climate change impact projects, many studies that investigate the usage of LAMs with higher spatial and temporal resolution in hydrological modelling generally do not consider long-term simulations. The previously cited studies include LAM applications of one month (Maussion et al., 2011), 3 months (Liu et al., 2011), 9 months (Pavelsky et al., 2011), and one year (Kunstmann and Stadler, 2005), respectively. Only Rögnvaldsson et al. (2007b) consider 16 years (using a coarser grid of 8 × 8 km) and the simulations of Mölg et al. (2012) cover 3 years.

In this study, two winter seasons with different meteorological conditions were selected for further investigations. This selection was restricted to the last decade, for which hourly observations are available. Unless otherwise stated, the following comments refer to Braunlage meteorological station (607 m a.s.l., Fig. 1). The winter season 2005/06 was colder than average. This holds especially for the period from January 2006 until March 2006. Moreover, in February and March 30 % above average precipitation was recorded. In the beginning of April, rapid snowmelt occurred due to significant rise in temperature accompanied by rain. The winter season 2010/11 on the other hand differs from average with respect to timing of accumulation and melt. In December 2010 the observed temperature was 5 K colder than

the average while precipitation was above average indicating ideal conditions for intensive snowfall. The highest daily snow depths in the second half of December observed at Clausthal for the period 1951–2011 were recorded in December 2010. In early January 2011 the weather changed and above average temperatures as well as rainfall were observed throughout the remaining winter season. In 2011, spring was exceptionally dry and warm. It is assumed that these differences in meteorological boundary conditions are sufficient to fulfil the prerequisites of a differential split sample test (Klemeš, 1986). With this type of test, different conditions for calibration and validation periods are presumed, and therefore, it “allows testing the ‘risky’ predictions of a model rather than the ‘safe’ ones” (Seibert, 2003). This approach allows to consider two different winter seasons at still reasonable computational costs.

2.3 Dynamical downscaling using WRF

In order to provide meteorological data fields with high resolution in space and time, a non-hydrostatic limited-area model (LAM) was applied. Non-hydrostatic LAMs are considered preferable for spatial resolutions of less than 10 km, as also suggested by Warner (2011). The freely available LAM Advanced Research WRF (Weather Research and Forecast modeling system, Skamarock et al., 2008) was chosen for this reason.

A multi-nesting approach setup of the model was established to provide a spatial resolution of 1.1 km for the study area (see Fig. 2). The first domain ($\Delta x = \Delta y = 30$ km) covers central Europe bounded by the North Sea and the Baltic Sea to the north and the Alps to the south. Due to further refinements of the grid resolution ($\Delta x = \Delta y = 1.1$ km), major characteristics of the topography of the study area are captured in the fourth domain.

The parameterization schemes of the model used in this study are listed in Tab. 1. A detailed description of the listed parameterization approaches can be found in the user’s guide (Wang et al., 2012) or in the cited literature. Convection parameterizations are only applied to domain 1 and 2. Moreover, four-dimensional data assimilation (FDDA, Stauffer and Seaman, 1990) was activated to keep the simulations close to the input data. This approach is proposed by Lo et al. (2008) although other approaches like the re-initialization are also common (Maussion et al., 2011; Hines et al., 2011).

Initial and boundary conditions were obtained from the NCEP FNL Operational Model Global Tropospheric Analyses dataset (ds083.2) using the WRF Preprocessing System (WPS) which is also described by Wang et al. (2012). This dataset contains major meteorological variables at the surface and mandatory vertical levels of the atmosphere on $1^\circ \times 1^\circ$ grids prepared operationally every six hours (NCEP, 2012a). Its temporal availability covers the range from July 1999 to present. To account for temperature changes in the North Sea and Baltic Sea, additional sea surface temperature data being available at daily grids with a spatial resolution of $0.5^\circ \times 0.5^\circ$ (RTG-SST dataset, Thiebaut et al., 2003; NCEP, 2012b) were also prepared using WPS.

Finally, hourly gridded meteorological fields with grid spacings of 1.1 km were simulated using WRF for both the winter season 2005/06 and 2010/11. Surface meteorological variables that were derived in this manner are: precipitation intensity, temperature (2 m), specific and relative humidity (2 m), wind speed (10 m), and shortwave as well as longwave radiation.

Depending on the computer hardware, WRF applications run 3 (24 cores) up to 20 days (4 cores) in order to simulate the meteorological fields of one single winter season for the domains shown in Fig. 2. A long-term WRF simulation, as would be desirable in this study, requires the application of a large-capacity computer causing relatively high computation costs. In order to avoid such costs and to show a pragmatic alternative for future users, who want to apply WRF downscaling but do not have access to large-scale computers, we selected only two winter seasons that represent different meteorological conditions as described in Sect. 2.2.

2.4 Snow models

Three energy balance snow models and a temperature-index model were tested at the point and at the catchment scale. All algorithms were originally available at the point scale. The following list accounts for the major characteristics of each model. The reader is referred to the cited literature for a detailed description. Because large parts of the Sieber catchment are covered by forests, the short descriptions emphasize the consideration of forest effects on the energy balance.

- 5 – The Utah Energy Balance Model Version 2.2 (Tarboton and Luce, 1996) is available as open source software and can be obtained via internet (Tarboton, 2012). Originally, the model was designed to carry out simulations for time steps of 1 and 6 h. Snow water equivalent (SWE) and energy content are the prognostic variables calculated for each time step using a predictor corrector approach. The model accounts for liquid water and incorporates a simple approach to simulate the ground heat flux. Among others, the most important parameters for calibrating the model are as follows (Tarboton and Luce, 1996): roughness length, thermal conductivity of snow, albedo of fresh snow, and another parameter that accounts for forest cover effects. The latter scales wind speed and solar irradiance. Snow accumulation is parametrized using two temperature related thresholds, which also enables the occurrence of mixed liquid and solid precipitation (mixed phase transition temperature range).
- 10
- 15 – ESCIMO (Energy balance Snow Cover Integrated MOdel, see, Strasser et al., 2002; Strasser and Marke, 2010) was originally developed for point-scale applications. The model was extended to the fully distributed hydro-climatological model AMUNDSEN (Alpine MULTiscale Numerical Distributed Simulation ENgine, Strasser et al., 2004, 2008), which includes detailed process descriptions, e.g. for topography-dependent radiative transfer, gravitational and wind-induced redistribution of snow, technical snow production (Hanzer et al., 2014) or runoff concentration. For both models, a detailed description of snow canopy interaction was added recently (Strasser et al., 2008, 2011). This approach includes subcanopy modifications to open-site meteorological conditions and incorporates a canopy interception model (Hedstrom and Pomeroy, 1998), which builds upon the scaling approach of Pomeroy and Schmidt (1993) and Pomeroy et al. (1998). The dependence of the leaf area index (LAI) on the interception capacity is accomplished according to Liston and Elder (2006). All these processes were also successfully tested at the catchment scale (Warscher et al., 2013). In this study, the point-scale model ESCIMO including the canopy model is used. The parameter set that is altered in order to calibrate the model includes albedo recession constants as well as the LAI.
- 20
- 25

ESCIMO includes a fixed wet-bulb temperature threshold of 2.0°C in order to separate rain and snow.

- Walter (2012) provides a spread sheet version of the algorithms described in Walter et al. (2005). The basic idea of their investigations was to develop a physically based alternative to the temperature index method without any additional data requirements. Regardless of its simplicity, the algorithm accounts for all relevant components of the energy balance as well as for liquid water storage and variable density. The original algorithm assumes daily time steps. In this study, all simulations were carried out for hourly time steps. Therefore, an adaptation of the original algorithm was necessary. Only the radiation balance has been modified in this study. Instead of using daily minimum and maximum temperature as proposed by Walter et al. (2005), time series of shortwave and longwave radiation are used. Several equations, such as the albedo recession, were adopted for hourly time steps. As does the Utah Energy Balance model, the snow model from Walter et al. (2005) accounts for forest canopy effects by providing one parameter to scale shortwave radiation. Furthermore, a temperature threshold for separating rain and snow as well as albedo recession constants can be altered in order to calibrate the model. All other processes like, e.g. turbulent fluxes, are originally parameterized using fixed values.
- For reasons of comparability and because of its popularity, the temperature-index approach was also considered. Forest effects on snowmelt can be taken into consideration by calibration of the degree-day value that is typically lower in forests compared to open sites. The degree-day factor is assumed to be independent of time. Snow is accumulated if the temperature falls below a given threshold temperature. Other enhancements, which are sometimes used for this approach, like, e.g., cold content, radiation input or liquid water storage are neglected in order to test the most basic version of this model.

The parameters for each snow model are listed in Tab. 2. In order to make the snow models available for catchment scale applications, they have been implemented into a hydrological modelling system. This step is described in the next section.

2.5 Hydrological modelling

All simulations at the catchment scale were carried out using the hydrological modelling system PANTA RHEI for the Sieber catchment. PANTA RHEI has been developed by the Department of Hydrology, Water Management and Water protection, Leichtweiss Institute for Hydraulic Engineering and Water Resources, University of Braunschweig, in corporation with the Institut für Wassermanagement IfW GmbH, Braunschweig, Germany (LWI-HYWAG and IfW, 2012). At present, the model is used for a wide range of tasks in national and international projects:

- Science: longterm water balance simulations and optimization of reservoir cascade operation for climate change impact studies (Meon and Gocht, 2012).
- Engineering practice: integrated flood protection concepts, hydrological design floods for hydraulic structures.
- Engineering practice: online flood prediction of medium to small catchments, e.g. by the Flood Early Warning Centre of the German Federal State of Lower Saxony.

Since PANTA RHEI features a graphical user interface (GUI) and geographic information system (GIS) data exchange capabilities, it is a 4th generation hydrological modelling system according to the classification of Refsgaard (1996). With respect to the level of sophistication of processes PANTA RHEI is a conceptual deterministic modelling system although physically based enhancements exist like the energy balance snow models described herein or physically based algorithms for evapotranspiration and the soil water balance (Förster et al., 2012; Kreye et al., 2012). It can also be characterized as a semi-distributed modelling system, subdividing the watershed into highly resolved sub-catchments and subsequently into hydrological response units (HRUs). The choice of the simulation time step depends on the task. Common settings range from minutes to one day. For flood simulations, one hour is a typical time step (LWI-HYWAG and IfW, 2012).

Observed precipitation is corrected with respect to systematic errors assuming correction factors for snow and rain. To provide meteorological data for each sub-catchment, a simple

distance related distribution method for point scale observations is applied. A constant lapse rate is assumed in order to account for altitudinal variations of temperature.

The vertical column of hydrological processes including snowmelt, interception, infiltration, and soil water are calculated for every HRU (Fig. 3). Runoff concentration (including several runoff components) and routing calculations are carried out for the entire sub-catchment presuming an aggregation of all associated HRUs. For this study we used the following configuration: the snow models described in Sect. 2.4 have been integrated into the modelling system as individual components, each representing a fully functioning snow model. The respective parameter accounting for forest cover effects, which is included in each of the models, is provided through lookup tables in PANTA RHEI in order to account for different stand densities in forested areas.

None of other hydrological process descriptions were changed throughout the study. Interception is calculated according to the Rutter et al. (1971) model. Infiltration and the soil water balance were simulated in a simplified way using a modified curve number approach which was adopted for continuous simulations (Riedel, 2004). Potential evapotranspiration is derived using the Penman–Monteith method (Monteith, 1965). For several land use classes respective vegetation parameters (e.g., LAI) have been provided in terms of lookup tables that were derived from literature review (e.g., Breuer et al., 2003). The calculation of actual evapotranspiration depends on the amount of intercepted water and soil water storage. For each HRU the runoff is subdivided into surface runoff, interflow, and base flow for each time step. A subdivision assuming two groundwater storages is also possible. Then these runoff components are aggregated to obtain the respective values for the superordinate sub-catchment. Routing is not only available for the river reaches of the sub-catchments but also for reservoirs, retention structures, culverts and other features.

A detailed model setup of the Sieber catchment for PANTA RHEI derived in an ongoing climate change research project was available for this study (Hölscher et al., 2012). The catchment was sub-divided into 73 sub-catchments which account for a mean area of 0.6 km². This sub-division into sub-catchments and HRUs was carried out based on digital elevation data, land use and soil information. Redirections of water due to channels associated to the

Upper Harz Water Management System were also considered in this context presuming mean annual and seasonal flow rates, which have been observed using gauging stations (Eggelsmann and Lange, 2011).

2.6 Model calibration

5 The calibration of the hydrological model and the snow models at both scales was carried out through maximizing the Nash-Sutcliffe model efficiency (see, e.g., Hall, 2001) manually. The hydrological model of the Sieber catchment was calibrated using long-term simulations based on daily meteorological data ranging from 1971 to 1991 (Nash-Sutcliffe model efficiency E for daily time step $E_{24h}=0.72$). In total 10 parameters were altered for the entire catchment
10 (one parameter for scaling of distributed degree-day values as well as a temperature threshold separating rain and snow, 3 parameters of the soil model, 3 recession constants for runoff concentration, and 2 parameters of the routing model). Subsequently, the model was validated using daily time series from 1991 to 2001 ($E_{24h}=0.79$). In order to adapt the model for
15 hourly time series of precipitation, a second calibration step was considered using hourly precipitation time series ranging from 2002 to 2008 ($E_{1h}=0.71$). This second step only included the adjustment of one parameter that separates infiltrated water and surface runoff in the model. Hence, it is assumed that the altered parameterization is suitable for simulations using hourly time steps.

The snow models were calibrated separately for the winter season 2005/06 (1 November 2005–1 May 2006), whereas the winter season 2010/11 (1 November 2010–1 April 2011)
20 is viewed as validation period. Nash-Sutcliffe model efficiencies are calculated for melt runoff at the point scale and streamflow at the catchment scale. As described in Sect. 2.4, the data requirement of models differ with respect to meteorological input. For the temperature-index approach only precipitation and temperature time series are mandatory, whereas the energy
25 balance approaches require radiation, humidity, and windspeed (mod. Walter approach: only shortwave and longwave radiation are additional requirements). Besides melt runoff only precipitation and temperature are recorded at Torfhaus. We therefore decided to prepare a uniform dataset, which was used to calibrate each of the snow models at both scales. This dataset

includes all relevant meteorological time series derived by WRF except for precipitation in order to avoid misleading configurations caused by precipitation uncertainty at small temporal scales. Observed precipitation time series is used instead.

All relevant parameters that were altered for each snow model in order to maximize the model efficiency with respect to melt runoff and streamflow, respectively, are summarized in Tab. 2. For point-scale applications none of the adjustable parameters of ESCIMO were altered at this stage. Calibrating the Utah Energy Balance model was most time-consuming since 6 parameters had to be altered. All parameters are within the range of values given by Tarboton and Luce (1996).

Since three quarters of the Sieber catchment are covered by forest, an adaptation to the point-scale parameter sets was necessary in order to calibrate each of the snow models, which are components of PANTA RHEI at the catchment scale. Thus, the calibration procedure for catchment-scale applications focused on the consideration of these effects. PANTA RHEI provides look-up tables including land-use dependent parameters that are relevant for the snow models including degree-day values, vegetation fraction, and LAI (among others). In order to calibrate the temperature-index model, a scaling factor for the HRU-based degree-day values was altered. Similarly, the land-use dependent vegetation fraction values needed to be adjusted for the modified Walter approach and the Utah Energy Balance Model using a scaling factor.

In contrast to all other tested snow models, ESCIMO includes a detailed model component to simulate snow processes in canopies (thus, the model is hereinafter referred to as ESCIMO + Canopy). This model utilizes the spatial distribution of LAI values for all areas covered by forests in the Sieber catchment. There is one adjustable parameter available that linearly scales distributed LAI values. For catchment-scale applications no adaptation of parameters was necessary in order to run ESCIMO + Canopy.

Table 3 summarizes the model efficiency achieved for each of the snow models at the point and the catchment scale, respectively. The model performance is generally better for the catchment scale compared to point scale considerations. Moreover, the range of model efficiencies is also smaller at the catchment scale.

3 Results and discussion

3.1 Snowmelt simulations using observed meteorological data

3.1.1 Point scale

As described earlier, only temperature and precipitation time series are available for Torfhaus meteorological station. Hence, this data only allows the application of the temperature-index approach. All simulations were carried out using a one hour time step for the entire winter season 1 November 2005–1 May 2006 as well as 1 November 2010–1 April 2011. Observed precipitation was corrected to account for systematic errors. Figure 4 shows SWE simulations as well as observed snow depth, which is recorded using an ultrasonic sensor. The course of SWE, as modelled using the temperature-index model, tracks the observed snow depth very well for both seasons. However, the shallow snow cover observed at in February and March 2011 is not captured by the simulation.

The corresponding time series of melt runoff for the same configuration are depicted in Fig. 5. In contrast to the SWE plot only the melt event is shown here. Every sub-plot shows the melt runoff recorded by the lysimeter. Cumulative runoff time series for both model and observation can easily be related to SWE and cumulative precipitation. In general, simulations match observed time series well, which holds also for diurnal features of melt runoff. However, some of the melt peaks in 2006 are overestimated, whereas one peak is not captured by the model. The total melt depth is simulated in accordance with observations. The model efficiency for the winter season 2005/06, which is calculated for an one hour time step, is $E=0.45$.

The model efficiency achieved for the winter season 2010/11 accounts for $E=0.75$. This value is better than the respective model efficiency of the winter season 2005/06 although the total melt depth is underestimated. These results emphasize the applicability of the previously calibrated temperature-index model using local observations at the point scale with respect to both SWE and melt runoff.

3.1.2 Catchment scale

According to the previous section, PANTA RHEI using the temperature-index approach was ran using observed meteorological time series for the winter season 2005/06. Since basin scale observations of SWE are not available the results are displayed for streamflow for the entire catchment only. Using the temperature index model at the catchment scale yields $E_{2006}=0.91$ for the winter season 2005/06 (Fig. 6). Modelled streamflow tracks the observed time series with high accuracy. However, the rain on snow event at the end of March to the beginning of April 2006 yielding $29.5 \text{ m}^3 \text{ s}^{-1}$ is underestimated by the model, which only accounts for $24 \text{ m}^3 \text{ s}^{-1}$.

The winter season 2010/11 is also simulated in accordance with observations ($E_{2011}=0.83$). The characteristics of the model performance is similar to the results obtained for the winter seasons 2005/06. Modelled and observed time series coincide well, which is expressed in the high model efficiency. However, the melt peak is also underestimated.

3.2 Snowmelt simulations using downscaled WRF data

3.2.1 Meteorological fields

To assess the accuracy of the downscaled WRF data, the meteorological data fields derived by WRF were compared to observations from meteorological stations. Figure 7 depicts the cumulative mean areal precipitation based on observations from the precipitation gauge network and the WRF simulations for the winter season 2005/06. The cumulative precipitation is derived by calculating the arithmetic mean of 19 stations and 19 corresponding grid points of the atmospheric model, respectively. These stations are situated in the Harz Mountains and its vicinity (see, Fig. 1). The temporal course coincides well with observations. However, the total sum of simulated precipitation only accounts for 81 % of the observed areal mean, which is not corrected with respect to systematic errors for this comparison. During the melt season, which falls at the end of March and the beginning of April 2006, simulated precipitation intensities are

substantially lower than observed intensities. The correlation coefficient of hourly intensities is $r = 0.52$.

Table 4 provides a statistical evaluation of hourly precipitation intensity for the stations that were included to calculate the areal mean. WRF underestimates the precipitation depth for the winter season 2005/06. The mismatch is greatest for Brocken meteorological station, which is located at the highest summit of the study area (1142 m a.s.l.). It is suspected that the complex topography is not resolved properly by the WRF model setup since the corresponding 1,1 km grid cell is located at 1008 m a.s.l. The Nash-Sutcliffe model efficiency E of most stations is below zero indicating a low model performance with respect to precipitation, even though the results are statistically reasonable, in particular at larger scales. For example, the statistical evaluation of the areal mean yields $E=0.14$. This value also increases if the temporal scale is shifted towards longer time steps (e.g., to 0.36 if a time step of six hours is specified). At the daily time scale a model efficiency of 0.52 can be achieved. The reasonably good match of the precipitation time series plotted in Fig. 7 is also reflected by the model efficiency of 0.80 calculated for a monthly time step.

The deviation between observed and simulated precipitation is obviously caused by uncertainties involved in the original analysis data and the downscaling process with regard to the heterogeneous orographic conditions of the small project region (upwind and downwind effects). It is not clear whether the generally lower WRF performance at the station locations in the lowlands is caused by lower model performance due to the nearby model domain boundary or due to weaker representation of hydrometeorological processes in this elevation band.

In general, temperature simulations match observations better than precipitation simulations do. The temperature time series of the WRF grid point that corresponds to the location of Torfhaus station was evaluated (Fig. 8). The simulated time series match observations at Torfhaus very closely, which is reflected in the high correlation of $r=0.93$. The model efficiency E amounts to 0.81. Nevertheless, the statistics reveal that the simulation is biased towards colder temperatures. The mean simulated temperature is 0.7 K lower than the mean value (-1.2°C) of the observed time series. This also becomes evident when considering the underestimation of maximum temperatures as shown in Fig. 8.

All other relevant meteorological variables were also well simulated. In accordance to Tab. 4 the statistical evaluation of all remaining variables that are relevant for hydrological modelling are summarized in Tab. 5. The model performance achieved for the variables strongly depends on the statistical characteristics of the meteorological variables. The model performance of meteorological variables that are subjected to fluctuations at small temporal scales is generally lower when compared to variables that do not show this feature. This holds especially for wind speed and for downwelling longwave radiation. The latter for instance is strongly influenced by the sequence of presence and absence of single clouds. However, mean and standard deviation of these variables are simulated in accordance with observations. Moreover, modelled temperature and humidity time series coincide very well when compared to observations. Corresponding plots of all these meteorological time series are shown by Förster (2013).

3.2.2 Point scale

In contrast to the previously described model setup, which involves observed meteorological time series, the full range of meteorological surface variables derived through dynamical downscaling enables the application of energy balance snow models. Thus, this model setting is not restricted to the temperature-index approach.

The calibrated snow models were applied using all meteorological variables derived through dynamical downscaling using WRF for the winter season 2005/06, which is shown in Fig. 9a for SWE and Fig. 10 for melt runoff. All snow models are in agreement with observations when modelled SWE is compared to observed snow depth (Fig. 9a). Accumulation as well as snow cover depletion are simulated very well with respect to the timing of these processes. However, it is not clear which model is most accurate in modelling maximum SWE.

Figure 10 shows the results for the melt event in spring of 2006 at Torfhaus (calibration period). While the timing of melt runoff including diurnal features is modelled well by all snow models, the peak runoff at 31 March 2006 is not captured. As explained earlier, this mismatch can be related to the fact that precipitation is underestimated by WRF throughout this period. There is still a remaining snow pack at the end of the melt event in all simulation runs, whereas the automated snow depth recordings indicate a complete ablation of the snow cover. In contrast,

melt runoff was still observed even though no precipitation was recorded during this period. Hence, the results from the simulation runs seem to be realistic.

The validation of the models was performed using data of the winter season 2010/11 (Fig. 11). Modelled runoff melt time series coincide well compared to the recorded lysimeter observations. Especially the Utah Energy Balance Model and ESCIMO closely match observations. However, the peak runoff is generally overestimated by all snow models. At the seasonal scale the snow cover dynamics are captured very well by all models. The courses of SWE are in accordance with the observed snow depth time series (Fig. 9b). When comparing the model efficiencies for each run, it is evident that the results for the validation period (2010/11) are generally better. However, the modified Walter approach seems to overestimate the SWE in February and March 2011.

Based on these results, the Utah Energy Balance Model seems to perform best at Torfhaus ($E_{2006}=0.54$, $E_{2011}=0.72$), whereas ESCIMO performs slightly better for the winter season 2005/06 at the same location ($E_{2006}=0.56$). In contrast, the modified Walter approach shows the lowest model performance ($E_{2006}=0.25$, $E_{2011}=0.44$) at this scale. The ranking of models should, however, be interpreted with caution because only two winter seasons were evaluated.

3.2.3 Catchment scale

According to the previous simulations at the point scale, all snow models were likewise applied using PANTA RHEI for the Sieber catchment. Figure 12 shows the results of each snow model according to the previous section but for the entire winter season 2005/06 including the accumulation period. All results plotted in Fig. 12 were derived using downscaled meteorological variables. The simulated streamflow tracks the observed streamflow reasonably well for all snow models. The rain on snow event at the end of March to the beginning of April yielding $29.5 \text{ m}^3 \text{ s}^{-1}$ is markable underestimated by the models. As already explained, this issue is related to the fact that WRF underestimates precipitation intensity during this period. However, the snowpack evolution is simulated in accordance with all models.

The results of the validation period (winter season 2010/11) are shown in Fig. 13. The tested snow models perform well compared to observations, although none of the models is able to

capture the first flood peak in November 2010. Possible explanations for this mismatch could include a lack in representativeness of simulated precipitation. However, all models track the observed time series very well during the melt period. The flood peak in January is captured by all models.

5 The assumption that the mismatch between modelled and observed snowmelt can at least partly be explained by biases in simulated precipitation is also supported by a study by Förster (2013). In this study, simulated precipitation has been substituted by observations with all the other meteorological variables still provided by WRF. While this approach might be useful in practice due to the fact that observed precipitation data are usually easier available than other
10 meteorological data, the combined application of observed and simulated meteorological input does not allow to fully assess the potential of WRF data for snowmelt modelling. It is therefore not further followed in the present study.

It is worth noting that the application of energy balance approaches leads to slightly improved model efficiency when compared to the temperature-index approach. For example, the model efficiencies achieved using the modified Walter approach, which performs best at catchment
15 scale, are $E_{2006}=0.76$ for the calibration period and $E_{2011}=0.70$ for the validation period. The corresponding model efficiencies of the temperature-index approach are $E_{2006}=0.51$ and $E_{2011}=0.66$, respectively.

The point and the catchment scale results reveal that the differences in model performance
20 are smaller for the catchment scale. At the catchment scale, many more hydrological processes have to be incorporated for simulation, rather than exclusively considering snowmelt processes at the point scale. Other hydrological processes become more dominant when scaling up from the point to the catchment scale. Typically, the relevance of snowmelt on floods decreases with an increase of scale.

25 When comparing results from both scales, it is interesting to note that the application of energy balance models enables reliable simulations of snowmelt processes. As explained in the previous section, model ranking should be interpreted with caution due to the fact that only two sites and two winter seasons were considered for this study.

4 Summary and conclusions

To investigate the effect of meteorological forcing and snow model complexity on hydrological simulations, two types of meteorological forcing data as well as four different snow models were tested for the Sieber catchment, Harz Mountains, Germany. Three snow models include a detailed description of the energy balance. Moreover, a temperature-index approach was also considered. All models were set up for point- and catchment-scale analyses. For the latter, the energy balance models were added to the hydrological modelling system PANTA RHEI, which already included the temperature-index approach.

Observed meteorological time series including precipitation and temperature were used as input for the temperature-index approach. Since other meteorological observations are hardly available, a dynamical downscaling approach using global atmospheric analysis data to force more complex energy balance snow models at different scales was also considered. In order to derive meteorological data fields we applied the Weather Research and Forecast model (WRF) driven by NCEP analysis data. The meteorological data fields including precipitation, temperature, humidity, wind speed, shortwave radiation, and longwave radiation were prepared as gridded datasets for hourly time steps.

The meteorological data fields were extensively evaluated using time series from the existing meteorological station network. The accuracy of precipitation simulations strongly depends on the considered scales. At the seasonal scale downscaled areal precipitation matches observations statistically reasonable, whereas at smaller spatial and temporal scales some events are not captured by the model. For all other meteorological variables relevant for hydrological modelling, the simulated time series match surface observations well.

The temperature-index model performs well at both scales if observed time series of temperature and precipitation are used as meteorological forcing. When observed time series are replaced by downscaled meteorological fields derived through WRF, the model performance of the temperature-index model decreases. However, WRF provides a full set of meteorological fields, which are physically consistent, even for a one hour time step. Thus, energy balance approaches can be applied successfully using this configuration of input data. The model

performance of the energy balance models is comparable to their respective values achieved for the temperature-index model. In some cases, energy balance approaches perform better than the temperature-index model.

Most uncertainty in modelled melt runoff and streamflow, respectively, arises from the mismatch in downscaled precipitation at small time scales although precipitation was modeled reliably at the seasonal scale. Even though deterministic precipitation simulations are generally seen as major challenge since modelling cloud microphysical processes is “founded upon less than perfect observational and theoretical base” (Stensrud, 2007), atmospheric models provide physically consistent meteorological fields, which holds also with respect to precipitation at larger temporal and spatial scales. These meteorological fields, in turn, enable the applicability of energy balance approaches for scenario simulations. Since this type of snow models are generally seen as more reliably for changing boundary conditions than temperature-index models (e.g., Charbonneau et al., 1981; Walter et al., 2005; Pomeroy et al., 2007; Singh et al., 2009; Barry and Gan, 2011; Warscher et al., 2013; Marke et al., 2014b), this study encourages the application of energy balance approaches for scenarios with changing boundary conditions.

Acknowledgements. The meteorological data as well as the hydrological model of the Sieber catchment were provided by the KliBiW research project (“Global climate change and impacts to water resources management in Lower Saxony”) which has been conducted in cooperation with the Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, NLWKN). This project was funded by the Ministry of the Environment of the German Federal State of Lower Saxony. The atmospheric analysis data are available online, provided by the National Centers for Environmental Prediction/National Weather Service/NOAA/US Department of Commerce (2000), updated daily, see “NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999”. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <http://dx.doi.org/10.5065/D6M043C6>, accessed 16 November 2012. Observed snow lysimeter time series were kindly made available by the Harzwasserwerke GmbH, Hildesheim. We wish to thank Larissa van der Laan for helping to improve the English of the final version of this manuscript. Last but not least, the authors would also like to thank three anonymous referees, Ryan Webb, and Bettina Schaeffli for their helpful comments and suggestions that significantly helped to improve the quality of the manuscript.

References

- Anderson, E. A.: Development and testing of snow pack energy balance equations, *Water Resour. Res.*, 4, 19–37, 1968.
- Anderson, E. A.: National Weather Service River Forecast System – Snow Accumulation and Ablation Model, in: NOAA Technical Memorandum, edited by: NOAA, vol. NWS HYDRO-17, National Weather Service, Silver Spring, 1973.
- Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J.: Mountain hydrology of the western United States, *Water Resour. Res.*, 42, W08432, doi:10.1029/2005WR004387, 2006.
- Barry, R. G. and Gan, T. Y.: *The Global Cryosphere: Past, Present, and Future*, Cambridge University Press, Cambridge, 2011.
- Bernhardt, M., Liston, G. E., Strasser, U., Zängl, G., and Schulz, K.: High resolution modelling of snow transport in complex terrain using downscaled MM5 wind fields, *The Cryosphere*, 4, 99–113, doi:10.5194/tc-4-99-2010, 2010.
- Beven, K. J.: *Rainfall-Runoff Modelling: The Primer*, Wiley, Chichester, 2001.
- Beven, K. J.: Causal models as multiple working hypotheses about environmental processes, *CR Geosci.*, 344, 77–88, 2012.
- Breuer, L., Eckhardt, K., and Frede, H.-G.: Plant parameter values for models in temperate climates, *Ecol. Model.*, 169, 237–293, 2003.
- Charbonneau, R., Lardeau, J.-P., and Obléd, C.: Problems of modelling a high mountainous drainage basin with predominant snow yields, *Hydrological Sciences Bulletin*, 26, 345–361, 1981.
- Chen, F. and Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system, Part I: Model implementation and sensitivity, *Mon. Weather Rev.*, 129, 569–585, 2001.
- Clark, M. P., Kavetski, D., and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resources Research*, 47, W09 301, 1–16, 2011.
- Dudhia, J.: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model, *J. Atmos. Sci.*, 46, 3077–3107, 1989.
- Eggelsmann, F. and Lange, A.: *Der Wasserhaushalt des Westharzes: Hydrologische Untersuchungen 1941 - 2010*, Harzwasserwerke GmbH, Abteilung Wasserwirtschaft, Hildesheim, 2011 (in German).
- El-Sadek, A., Bleiweiss, M., Shukla, M., Guldán, S., and Fernald, A.: Alternative climate data sources for distributed hydrological modelling on a daily time step, *Hydrol. Process.*, 25, 1542–1557, 2011.

- Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.-J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, O., Pyles, R. D., Schlosser, A., Shmakin, A. B., Smirnova, T. G., Strasser, U., Verseghy, D., Yamazaki, T., and Yang, Z.-L.: Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project), *Ann. Glaciol.*, 38, 150–158, 2004.
- 5 Förster, K.: Detaillierte Nachbildung von Schneeprozessen in der hydrologischen Modellierung, Ph.D. thesis, Technische Universität Braunschweig, Braunschweig, available at: <http://www.digibib.tu-bs.de/?docid=00052554>, last access: 7 April 2014, 2013 (in German).
- Förster, K., Gelleszun, M., and Meon, G.: A weather dependent approach to estimate the annual course of vegetation parameters for water balance simulations on the meso- and macroscale, *Adv. Geosci.*, 10 32, 15–21, doi:10.5194/adgeo-32-15-2012, 2012.
- Franz, K. J.: Characterization of the Comparative Skill of Conceptual and Physically-Based Snow Models for Streamflow Prediction, Ph.D. thesis, University of California, Irvine, 2006.
- Franz, K. J., Hogue, T. S., and Soroshian, S.: Operational snow modeling: addressing the challenges of an energy balance model for National Weather Service forecasts, *J. Hydrol.*, 360, 48–66, 2008.
- 15 Fuka, D. R., Easton, Z. M., Brooks, E. S., Boll, J., Steenhuis, T. S., and Walter, M. T.: A simple process-based snowmelt routine to model spatially distributed snow depth and snowmelt in the SWAT Model, *J. Am. Water Resour. As.*, 48, 1151–1161, 2012.
- Giorgi, F.: Regional climate modeling: status and perspectives, *J. Phys. IV*, 139, 101–118, 2006.
- 20 Hall, M. J.: How well does your model fit the data?, *J. Hydroinform.*, 3, 49–55, 2001.
- Hanzer, F., Marke, T., and Strasser, U.: Distributed, explicit modeling of technical snow production for a ski area in the Schladming Region (Austrian Alps), *Cold Reg. Sci. Technol.*, in press, 2014.
- Hedstrom, N. R. and Pomeroy, J. W.: Measurements and modelling of snow interception in the boreal forest, *Hydrol. Process.*, 12, 1611–1625, 1998.
- 25 Hines, K. M., Bromwich, D. H., Bai, L.-S., Barlage, M., and Slater, A. G.: Development and testing of Polar WRF, Part III: Arctic Land*, *J. Climate*, 24, 26–48, 2011.
- Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, *J. Glaciol.*, 45, 101–111, 1999.
- Hölscher, J., Petry, U., Bertram, M., Anhalt, M., Schmidtke, S., Haberlandt, U., Müller, H., van der Heijden, S., Berndt, C., Verworn, A., Wallner, M., Belli, A., Dietrich, J., Meon, G., Förster, K., Gelleszun, M., Riedel, G., Lange, A., and Eggelsmann, F.: Globaler Klimawandel: Wasserwirtschaftliche Folgenabschätzung für das Binnenland, vol. 33 of *Oberirdische Gewässer, NLWKN, Norden*, 2012 (in German).

- Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, *Mon. Weather Rev.*, 134, 2318–2341, 2006.
- Jost, G., Dan Moore, R., Smith, R., and Gluns, D. R.: Distributed temperature-index snowmelt modelling for forested catchments, *J. Hydrol.*, 420–421, 87–101, 2012.
- 5 Kain, J. S.: The Kain–Fritsch convective parameterization: an update, *J. Appl. Meteorol.*, 43, 170–181, 2004.
- Klemeš, V.: Operational testing of hydrological simulation models, *Hydrolog. Sci. J.*, 31, 13–24, 1986.
- Kreye, P., Gelleszun, M., and Meon, G.: Ein landnutzungssensitives Bodenmodell für die meso- bis makroskalige Wasserhaushaltsmodellierung, in: *Wasser ohne Grenzen – Beiträge zum Tag der Hydrologie am 22./23. März 2012 an der Albert-Ludwigs-Universität Freiburg*, edited by: Weiler, M., vol. 31.12 of *Forum für Hydrologie und Wasserbewirtschaftung*, Fachgemeinschaft Hydrologische Wissenschaften in der DWA, Hennef, 25–30, 2012 (in German).
- 10 Kuchment, L. S. and Gelfan, A. N.: The determination of the snowmelt rate and the meltwater outflow from a snowpack for modelling river runoff generation, *J. Hydrol.*, 179, 23–36, 1996.
- 15 Kunstmann, H. and Stadler, C.: High resolution distributed atmospheric-hydrological modelling for Alpine catchments, *J. Hydrol.*, 314, 105–124, 2005.
- Liston, G. E. and Elder, K.: A Distributed Snow-Evolution Modeling System (SnowModel), *J. Hydrometeorol.*, 7, 1259–1276, 2006.
- Liu, C., Ikeda, K., Thompson, G., Rasmussen, R. M., and Dudhia, J.: High-resolution simulations of wintertime precipitation in the Colorado Headwaters Region: sensitivity to physics parameterizations, *Mon. Weather Rev.*, 139, 3533–3553, 2011.
- 20 Lo, J. C.-F., Yang, Z.-L., and Pielke, R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, *J. Geophys. Res.*, 113, D09112, doi:10.1029/2007JD009216, 2008.
- 25 LWI-HYWAG and IfW: PANTA RHEI User Manual (unpublished), Department of Hydrology, Water Management and Water Protection (HYWAG), Leichtweiss Institute (LWI) for Hydraulic Engineering and Water Resources, University of Braunschweig and Institut für Wassermanagement IfW GmbH, Braunschweig, 2012.
- Marke, T., Mauser, W., Pfeiffer, A., Zängl, G., and Jacob, D.: The effect of downscaling on river runoff modeling: a hydrological case study in the Upper Danube Watershed, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 6331–6384, doi:10.5194/hessd-8-6331-2011, 2011.
- 30 Marke, T., Mauser, W., Pfeiffer, A., Zängl, G., Jacob, D., and Strasser, U.: Application of a hydrometeorological model chain to investigate the effect of global boundaries and downscaling

- on simulated river discharge, *Environ. Earth Sci.*, 71, 4849–4868, doi:10.1007/s12665-013-2876-z, 2014a.
- Marke, T., Strasser, U., Hanzer, F., Wilcke, R. A. I., and Gobiet, A.: Scenarios of future snow conditions in Styria (Austrian Alps), *J. Hydrometeorol.*, in press, 2014b.
- 5 Maussion, F., Scherer, D., Finkelnburg, R., Richters, J., Yang, W., and Yao, T.: WRF simulation of a precipitation event over the Tibetan Plateau, China – an assessment using remote sensing and ground observations, *Hydrol. Earth Syst. Sci.*, 15, 1795–1817, doi:10.5194/hess-15-1795-2011, 2011.
- Meon, G. and Gocht, M.: Minimization of adverse climate change impacts on multi-dam systems by optimized operation, in: *Proceedings of the International Symposium on Dams for a Changing World, 80th Annual Meeting and 24th Congress of ICOLD, 5 June 2012, 1.33–1.38, Kyoto, 2012.*
- 10 Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: implications for climate impact analyses, *Water Resour. Res.*, 47, W02531, 1–17, 2011.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.*, 102, 16663–16682, 1997.
- 15 Mölg, T., Maussion, F., Yang, W., and Scherer, D.: The footprint of Asian monsoon dynamics in the mass and energy balance of a Tibetan glacier, *The Cryosphere*, 6, 1445–1461, doi:10.5194/tc-6-1445-2012, 2012.
- Monteith, J. L.: Evaporation and environment, in: *The State and Movement of Water in Living Organisms*, edited by: Fogg, G. E., vol. 19 of *Symposia of the Society for Experimental Biology*, Cambridge University Press, Cambridge, 205–234, 1965.
- 20 Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes, *Mon. Weather Rev.*, 137, 991–1007, 2009.
- 25 NCEP: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999: ds083.2, doi:10.5065/D6M043C6, 2012a.
- NCEP: Real-time, global, sea surface temperature (RTG_SST) analysis, available at: <http://polar.ncep.noaa.gov/sst/oper/>, last access: 7 April 2014, 2012b.
- Pavelsky, T. M., Kapnick, S., and Hall, A.: Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California, *J. Geophys. Res.*, 30 116, D16115, 1–18, 2011.

- Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., and Corripio, J.: An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland, *J. Glaciol.*, 51, 573–587, 2005.
- 5 Pomeroy, J. W. and Schmidt, R. A.: The use of fractal geometry in modeling intercepted snow accumulation and sublimation, in: *Proceedings of the 50th Eastern Snow Conference, 61st Western Snow Conference*, Ville de Québec, 1–10, 1993.
- Pomeroy, J. W., Parviainen, J., Hedstrom, N. R., and Gray, D. M.: Coupled modelling of forest snow interception and sublimation, *Hydrol. Process.*, 12, 2317–2337, 1998.
- 10 Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J., and Carey, S. K.: The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence, *Hydrol. Process.*, 21, 2650–2667, 2007.
- Rango, A. and Martinec, J.: Revisiting the degree-day method for snowmelt computations, *J. Am. Water Resour. As.*, 31, 657–669, 1995.
- 15 Refsgaard, J. C.: Terminology, Modelling Protocol and Classification of Hydrological Model Codes, in: *Distributed Hydrological Modelling*, edited by: Abbott, M. B. and Refsgaard, J. C., vol. 22 of *Water Science Technology Library*, Kluwer Acad. Publ., Dordrecht, 17–39, 1996.
- Riedel, G.: Ein hydrologisches Modell für tidebeeinflusste Flussgebiete, Ph.D. thesis, Technische Universität Braunschweig, Braunschweig, available at: <http://www.digibib.tu-bs.de/?docid=00001644>, last access: 7 April 2014, 2004 (in German).
- 20 Rögnvaldsson, Ó., Bao, J.-W., and Ólafsson, H.: Sensitivity simulations of orographic precipitation with MM5 and comparison with observations in Iceland during the Reykjanes Experiment, *Meteorol. Z.*, 16, 87–98, 2007a.
- Rögnvaldsson, Ó., Jónsdóttir, J. F., and Ólafsson, H.: Numerical simulations of precipitation in the complex terrain of Iceland – comparison with glaciological and hydrological data, *Meteorol. Z.*, 16, 71–85, 2007b.
- 25 Rummukainen, M.: State-of-the-art with regional climate models, *WIREs Clim. Change*, 1, 82–96, doi:10.1002/wcc.8, 2010.
- Rutter, A. J., Kershaw, K. A., Robins, P. C., and Morton, A. J.: A predictive model of rainfall interception in forests, 1. Derivation of the model from observations in a plantation of Corsican pine, *Agr. Meteorol.*, 9, 367–384, 1971.
- 30 Rutter, N., Essery, R., Pomeroy, J. W., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A. N., Goodbody, A., Gusev, Y., Gustafsson, D., Hellström, R., Hirabayashi, Y., Hirota, T., Jonas, T.,

- Koren, V., Kuragina, A., Lettenmaier, D. P., Li, W.-P., Luce, C., Martin, E., Nasonova, O., Pumpanen, J., Pyles, R. D., Samuelsson, P., Sandells, M., Schädler, G., Shmakin, A. B., Smirnova, T. G., Stähli, M., Stöckli, R., Strasser, U., Su, H., Suzuki, K., Takata, K., Tanaka, K., Thompson, E., Vesala, T., Viterbo, P., Wiltshire, A., Xia, K., Xue, Y., and Yamazaki, T.: Evaluation of forest snow processes models (SnowMIP2), *J. Geophys. Res.*, 114, D06111, doi:10.1029/2008JD011063, 2009.
- Seibert, J.: Conceptual Runoff Models – Fiction or Representation of Reality?, Ph.D. thesis, Uppsala University, Uppsala, Sweden, 1999.
- Seibert, J.: Reliability of model predictions outside calibration conditions: paper presented at the Nordic Hydrological Conference (Røros, Norway, 4–7 August 2002), *Nord. Hydrol.*, 34, 477–492, 2003.
- Singh, P. R., Gan, T. Y., and Gobena, A. K.: Evaluating a hierarchy of snowmelt models at a watershed in the Canadian Prairies, *J. Geophys. Res.*, 114, D04109, doi:10.1029/2008JD010597, 2009.
- Sivapalan, M., Takeuchi, K., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., Zehe, E., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., and O’Connell, P. E.: IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences, *Hydrolog. Sci. J.*, 48, 857–880, 2003.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, vol. NCAR/TN–475 + STR of NCAR Technical Note, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, 2008.
- Stauffer, D. R. and Seaman, N. L.: Use of four-dimensional data assimilation in a limited-area mesoscale model, Part I: Experiments with synoptic-scale data, *Mon. Weather Rev.*, 118, 1250–1277, 1990.
- Stensrud, D.: Parameterization schemes: Keys to understanding numerical weather prediction models, Cambridge University Press, New York, 2007.
- Strasser, U. and Marke, T.: *ESCIMO.spread* – a spreadsheet-based point snow surface energy balance model to calculate hourly snow water equivalent and melt rates for historical and changing climate conditions, *Geosci. Model Dev.*, 3, 643–652, doi:10.5194/gmd-3-643-2010, 2010.
- Strasser, U., Etchevers, P., and Lejeune, Y.: Inter-comparison of two snow models with different complexity using data from an Alpine site: selected paper from EGS General Assembly, Nice, April 2000 (Symposium OA36), *Nord. Hydrol.*, 33, 15–26, 2002.
- Strasser, U., Corripio, J., Pellicciotti, F., Burlando, P., Brock, B., and Funk, M.: Spatial and temporal variability of meteorological variables at Haut Glacier d’Arolla (Switzerland) during

- the ablation season 2001: Measurements and simulations, *J. Geophys. Res.*, 109, D03103, doi:10.1029/2003JD003973, 2004.
- Strasser, U., Bernhardt, M., Weber, M., Liston, G. E., and Mauser, W.: Is snow sublimation important in the alpine water balance?, *The Cryosphere*, 2, 53–66, doi:10.5194/tc-2-53-2008, 2008.
- 5 Strasser, U., Warscher, M., and Liston, G. E.: Modeling snow-canopy processes on an idealized mountain, *J. Hydrometeorol.*, 12, 663–677, 2011.
- Tarboton, D. G.: UEB Snow Model, available at: <http://www.neng.usu.edu/cee/faculty/dtarb/snow/snow.html>, last access: 7 April 2014, 2012.
- Tarboton, D. G. and Luce, C.: Utah Energy Balance Snow Accumulation and Melt Model (UEB):
10 Computer Model Technical Description and Users Guide, Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, 1996.
- Thiebaut, J., Rogers, E., Wang, W., and Katz, B.: A new high-resolution blended real-time global sea surface temperature analysis, *B. Am. Meteorol. Soc.*, 84, 645–656, 2003.
- Walter, M. T.: Snowmelt Model, available at: <http://www.hydrology.bee.cornell.edu/Papers/WalterSnow.xls>, last access: 7 April 2014, 2012.
- 15 Walter, M. T., Brooks, E. S., McCool, D. K., King, L. G., Molnau, M., and Boll, J.: Process-based snowmelt modeling: does it require more input data than temperature-index modeling?, *J. Hydrol.*, 300, 65–75, 2005.
- Wang, W., Bruyère, C., Duda, M., Dudhia, J., Gill, D., Kavulich, M., Keene, K., Lin, H.-C.,
20 Michalakes, J., Rizvi, S., Zhang, X., Beezley, J. D., Coen, J. L., Mandel, J., Chuang, H.-Y., McKee, N., Slovacek, T., and Wolff, J.: User’s Guide for the Advanced Research WRF (ARW) Modeling System Version 3.4, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Reserach, Boulder, 2012.
- Warner, T. T.: Numerical Weather and Climate Prediction, Cambridge University Press, Cambridge, 2011.
- 25 Warscher, M., Strasser, U., Kraller, G., Marke, T., Franz, H., and Kunstmann, H.: Performance of complex snow cover descriptions in a distributed hydrological model system: a case study for the high Alpine terrain of the Berchtesgaden Alps, *Water Resour. Res.*, 49, 2619–2637, 2013.

Table 1. WRF physics parameterization setup. Brief descriptions of the physics parameterizations are also given by Wang et al. (2012) and Skamarock et al. (2008).

Processes	Parameterization
Cloud microphysics	Morrison et al. (2009) two-moment bulk microphysics
Convection	Kain–Fritsch scheme (Kain, 2004)
Boundary layer	Yonsei University scheme (Hong et al., 2006)
Land surface processes	Noah land surface model (LSM) (Chen and Dudhia, 2001)
Shortwave radiation	Dudhia (1989) scheme
Longwave radiation	Rapid Radiative Transfer Model (Mlawer et al., 1997)

Table 2. Model parameters that were altered throughout the calibration process for both the point and the catchment scale. Please note that the catchment scale parameters refer to the entire Sieber catchment upstream of Pionierbrücke gauging station.

Snow model	Parameter	Point scale	Catchment scale
Temperature-Index	temperature separating rain and snow [$^{\circ}\text{C}$]	1	1.5
	degree-day value [$\text{mm d}^{-1} \text{K}^{-1}$]	6.6	-
	degree-day scaling factor [-]	-	2
mod. Walter et al.	albedo recession, $T < 0$ [d^{-1}]	0.05	0.1
	albedo recession, $T \geq 0$ [d^{-1}]	0.12	0.3
	temperature separating rain and snow [$^{\circ}\text{C}$]	1	2
Utah Energy Balance	forest cover fraction scaling factor [-]	-	0.75
	temperature separating rain and snow [$^{\circ}\text{C}$]	1	2
	mixed phase transition temp. range [K]	1	2
	snow surface aerodynamic roughness [mm]	0.3	0.5
	snow surface thermal conductance [m h^{-1}]	0.02	0.02
	new snow visible albedo [-]	0.9	0.7
	forest cover fraction scaling factor [-]	-	0.6
ESCIMO (+ Canopy)	albedo recession, $T < 0$ [d^{-1}]	0.05	0.05
	albedo recession, $T \geq 0$ [d^{-1}]	0.12	0.12
	leaf area index scaling factor	-	1

Table 3. Nash-Sutcliffe model efficiencies E derived through calibration at the point and the catchment scale, respectively. Statistics refer to the winter season 2005/06.

	Point scale	Catchment scale
Temperature-Index	0.49	0.85
mod. Walter et al.	0.32	0.90
Utah Energy Balance	0.71	0.82
ESCIMO (+Canopy)	0.40	0.84

Table 4. Comparison of observed and simulated precipitation time series for each station. This table includes the elevation of each station (z), simulated and observed precipitation depth for the winter season 2005/06 (P_s and P_o), the corresponding standard deviation of simulated and observed precipitation intensity ($\tilde{\sigma}_{P_s}$ and $\tilde{\sigma}_{P_o}$), correlation between simulated and observed precipitation (r), root mean square error (RMSE) and the model efficiency (E). The temporal resolution of this evaluation is $\Delta t=1$ h, which is available for 18 of 19 stations (performance measures for Brocken are excluded since the temporal resolution of these station recordings is $\Delta t=6$ h). Observed values are not corrected for this comparison.

station	z [m]	P_s [mm]	P_o [mm]	$\tilde{\sigma}_{P_s}$ [$\frac{\text{mm}}{\text{h}}$]	$\tilde{\sigma}_{P_o}$ [$\frac{\text{mm}}{\text{h}}$]	RMSE [$\frac{\text{mm}}{\text{h}}$]	r [-]	E [-]
Altenau	420	531.4	649.2	0.43	0.43	0.48	0.38	-0.25
Braunlage	607	619.6	664.1	0.51	0.49	0.53	0.45	-0.15
Brocken	1142	608.4	1275.6	-	-	-	-	-
Bühne-Rimbeck	100	186.6	191.5	0.21	0.22	0.25	0.32	-0.31
Clausthal	561	479.6	636.8	0.39	0.46	0.49	0.35	-0.14
Eckertalsperre	505	451.7	508.6	0.35	0.52	0.52	0.34	-0.01
Granetalsperre	311	316.0	308.7	0.28	0.28	0.33	0.32	-0.37
Innerstetalsperre	230	344.1	354.4	0.33	0.29	0.38	0.26	-0.70
Liebenburg-Othfresen	187	207.6	233.8	0.21	0.22	0.23	0.43	-0.08
Odertalsperre	335	414.7	633.7	0.35	0.54	0.54	0.33	0.01
Riefensbeek	345	436.7	634.9	0.36	0.46	0.46	0.37	-0.06
Schierke	609	632.5	564.9	0.48	0.42	0.46	0.49	-0.21
Schulenberg	470	541.3	619.8	0.44	0.46	0.50	0.38	-0.19
Seesen	186	249.4	301.5	0.25	0.27	0.31	0.29	-0.30
Sieber (Pionierbrücke)	340	509.6	602.4	0.43	0.60	0.60	0.35	-0.01
Sösetalsperre	275	295.8	533.0	0.27	0.41	0.41	0.37	0.04
Stiege	494	387.6	303.9	0.38	0.30	0.39	0.34	-0.71
Torfhaus	805	578.2	626.5	0.43	0.47	0.49	0.41	-0.09
Wernigerode	234	238.3	227.1	0.26	0.25	0.29	0.35	-0.37
areal mean	429	422.6	519.5	0.28	0.32	0.30	0.52	0.14

Table 5. Comparison of observed and simulated time series for the most relevant meteorological variables. This table includes simulated and observed mean of each variable for the winter season 2005/06 (\bar{x}_s and \bar{x}_o), the corresponding standard deviation ($\tilde{\sigma}_{x_s}$ and $\tilde{\sigma}_{x_o}$), root mean square error (RMSE), correlation (r), and the model efficiency (E). The superscripts denote the station for which the comparison was carried out. Station 1: Braunlage (607 m a.s.l., WRF domain 4), station 2: Braunschweig (81 m a.s.l., WRF domain 3). The time step of this evaluation is $\Delta t=1$ h.

	\bar{x}_s	\bar{x}_o	$\tilde{\sigma}_{x_s}$	$\tilde{\sigma}_{x_o}$	RMSE	r	E
	[unit of respective input time series is used]					[-]	[-]
Precipitation intensity ¹⁾ [mm h ⁻¹]	0.14	0.15	0.51	0.49	0.53	0.45	-0.15
Temperature ¹⁾ [K]	271.59	273.06	5.49	5.07	2.42	0.94	0.77
Specific humidity ¹⁾ [g kg ⁻¹]	3.56	3.86	1.42	1.41	0.55	0.95	0.85
Wind speed ¹⁾ [m s ⁻¹]	3.99	3.02	2.23	1.55	1.89	0.69	-0.47
Shortwave radiation ²⁾ [W m ⁻²]	56.53	61.22	138.72	124.09	79.75	0.82	0.59
Longwave radiation ²⁾ [W m ⁻²]	289.58	287.53	40.80	43.13	35.64	0.64	0.32

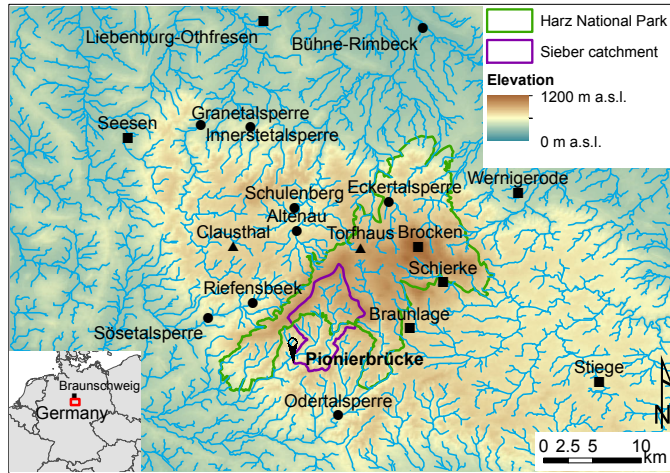


Figure 1. Map of the study area including the Sieber catchment (watershed boundary is magenta-coloured) located in the Harz Mountains, northern Germany. The station network shown on the map provides hourly time series for precipitation only (circles), precipitation and temperature (triangles), and additional observations (squares). The Harz National Park boundary is provided by OpenStreetMap, © OpenStreetMap contributors (<http://www.openstreetmap.org/copyright>).

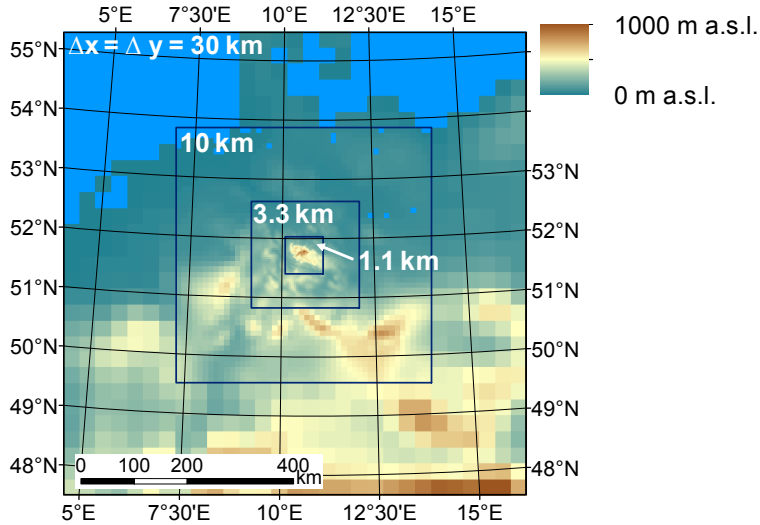


Figure 2. WRF domains of the study area. The resolution of the first (outer) domain is 30 km, the fourth and final (inner) domain's resolution is 1.1 km, respectively.

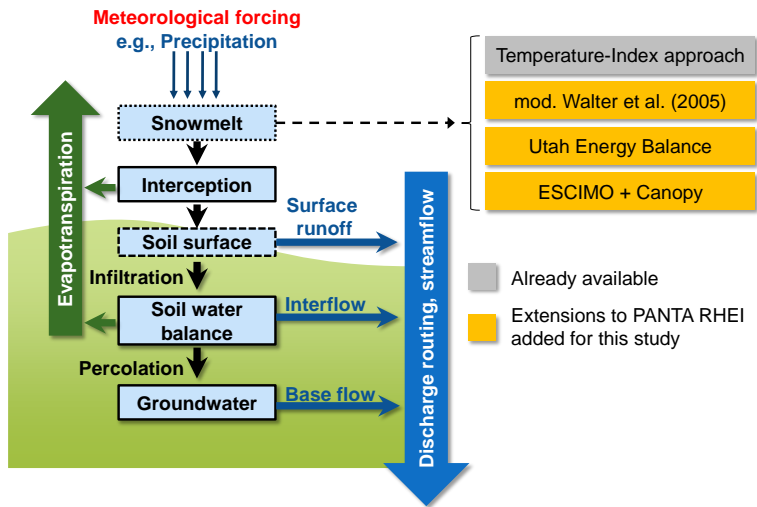


Figure 3. Conceptualization of hydrological processes in the hydrological modelling system PANTA RHEI (adapted from LWI-HYWAG and IfW, 2012). According to Sect. 2.4, the snowmelt process can be simulated by four different models.

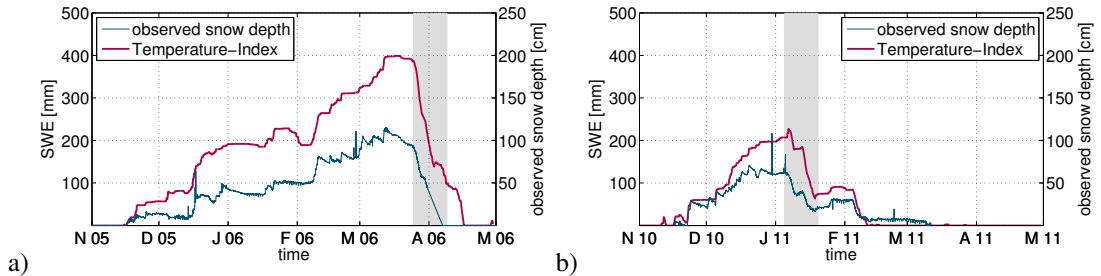


Figure 4. Comparison of simulated SWE and observed snow depth time series at Torfhaus meteorological station: **(a)** Winter season 2005/06. **(b)** Winter season 2010/11. Both simulations were performed using observed meteorological time series. The shaded area refers to the corresponding melt period (see, Fig. 5).

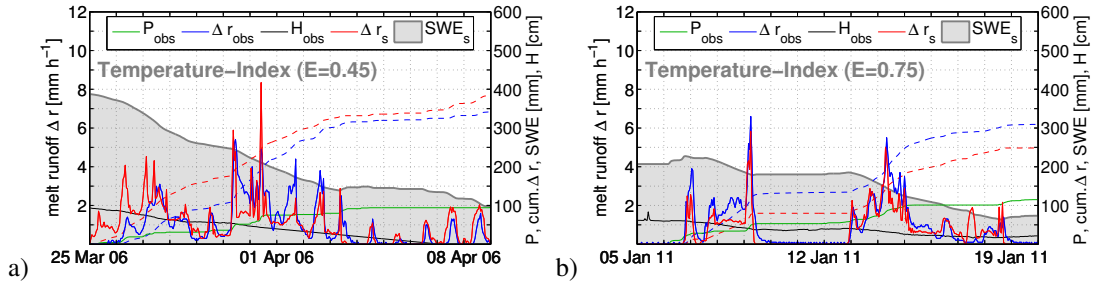


Figure 5. Snowmelt simulations at the point scale using observed meteorological time series at Torfhaus meteorological station: **(a)** Winter season 2005/06. **(b)** Winter season 2010/11. For every simulation run observed precipitation depth (P_{obs}), observed runoff (Δr_{obs} , blue line) including observed runoff depth (blue dashed line), observed snow depth (H_{obs}), simulated runoff (Δr_{s} , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_{s}) are plotted. Model efficiencies E are also given for each run.

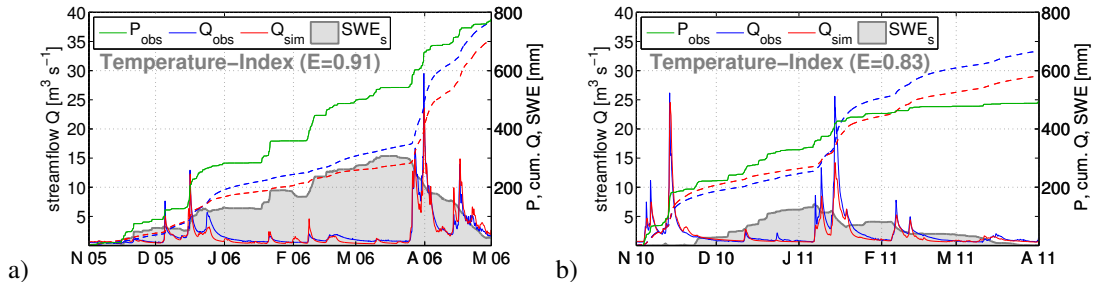


Figure 6. Streamflow simulations at the catchment scale using observed meteorological time series: **(a)** Winter season 2005/06. **(b)** Winter season 2010/11. For every simulation run observed precipitation depth (P_{obs}), observed streamflow (Q_{obs} , blue line) including observed runoff depth (blue dashed line), simulated streamflow (Q_{s} , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_{s}) are plotted. Model efficiencies E are also given for each run.

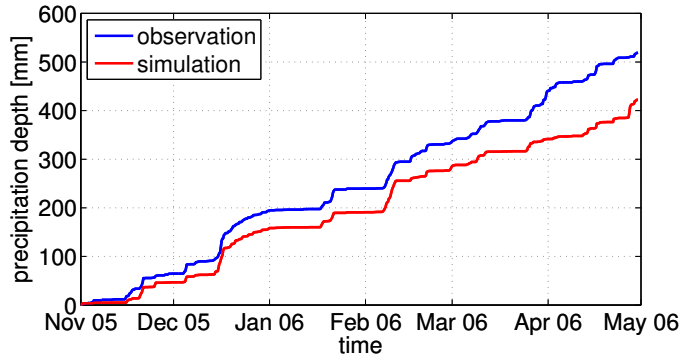


Figure 7. Comparison of observed and simulated mean areal precipitation depth. The cumulative precipitation depth is derived by calculating the arithmetic mean of 19 stations and 19 corresponding grid points of the limited-area model, respectively. At this stage no correction approach was applied to the observed time series, e.g. to correct wind induced errors. Please refer to Tab. 4 for further details including performance measures.

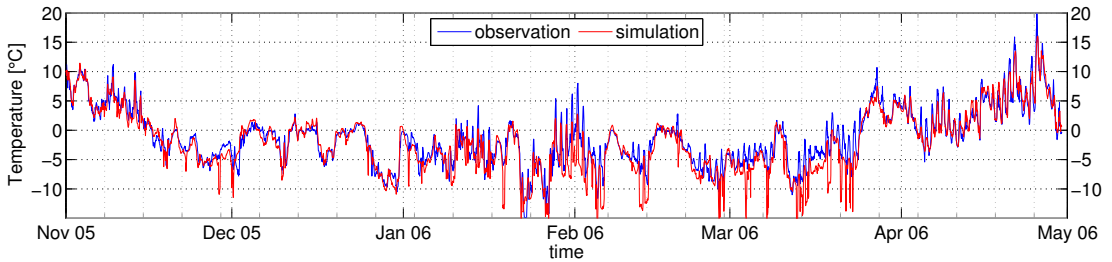


Figure 8. Comparison of simulated and observed temperature time series at Torfhaus meteorological station.

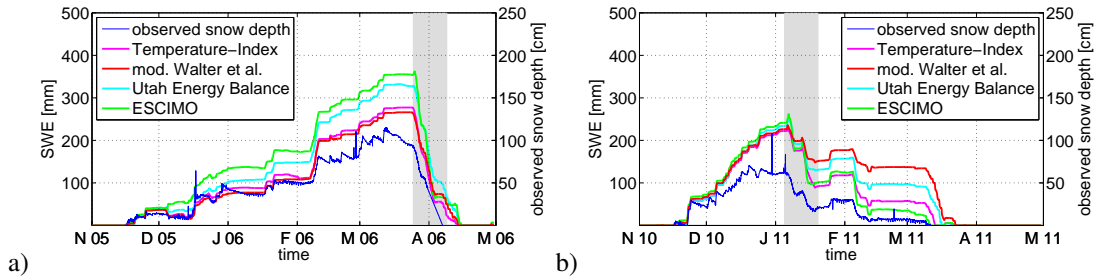


Figure 9. Comparison of simulated SWE and observed snow depth time series at Torfhaus meteorological station. (a) Winter season 2005/06. (b) Winter season 2010/11. This plot includes time series of SWE for all tested snow models using downscaled WRF meteorological input.

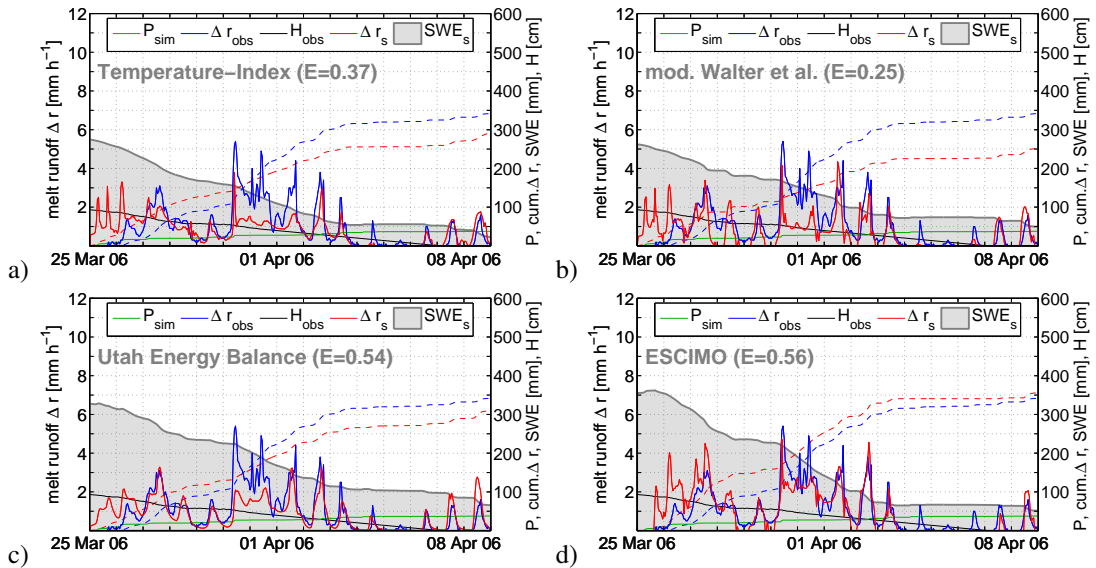


Figure 10. Snowmelt simulations at the point scale using downscaled meteorological time series for the winter season 2005/06. Statistics refer to the period 25 March–9 April 2006. (a) Temperature-Index approach, (b) modified Walter et al. (2005) approach, (c) Utah Energy Balance model, and (d) ESCIMO. For every simulation run downscaled precipitation depth (P_{sim}), observed runoff (Δr_{obs} , blue line) including observed runoff depth (blue dashed line), observed snow depth (H_{obs}), simulated runoff (Δr_{s} , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_{s}) are plotted. Model efficiencies E are also given for each run.

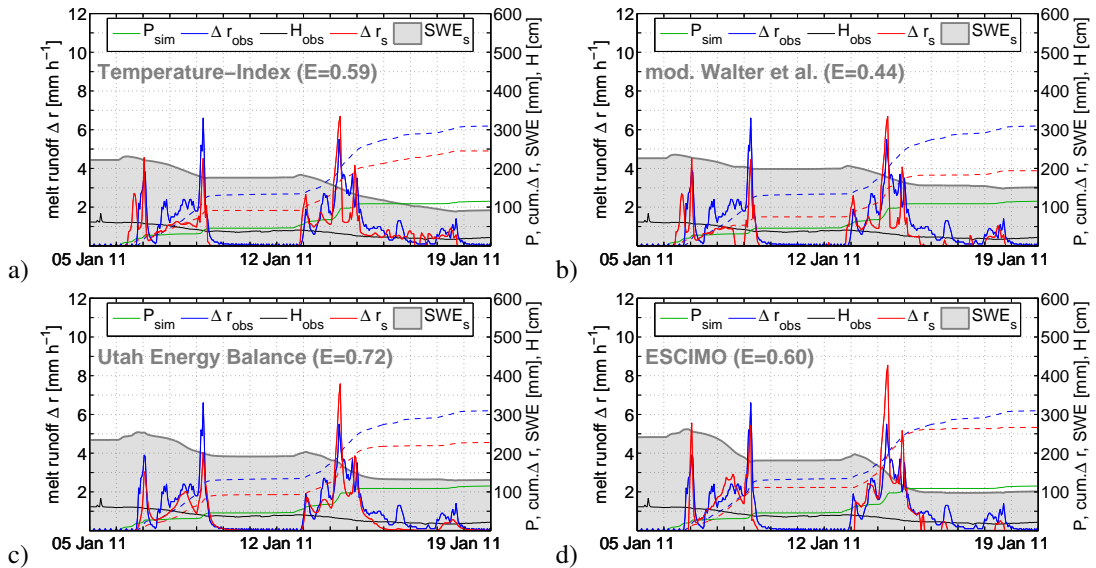


Figure 11. Snowmelt simulations at the point scale using downscaled meteorological time series for the winter season 2010/11. Statistics refer to the period 5–20 January 2011. **(a)** Temperature-Index approach, **(b)** modified Walter et al. (2005) approach, **(c)** Utah Energy Balance model, and **(d)** ESCIMO. For every simulation run downscaled precipitation depth (P_{sim}), observed runoff (Δr_{obs} , blue line) including observed runoff depth (blue dashed line), observed snow depth (H_{obs}), simulated runoff (Δr_{s} , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_{s}) are plotted. Model efficiencies E are also given for each run.

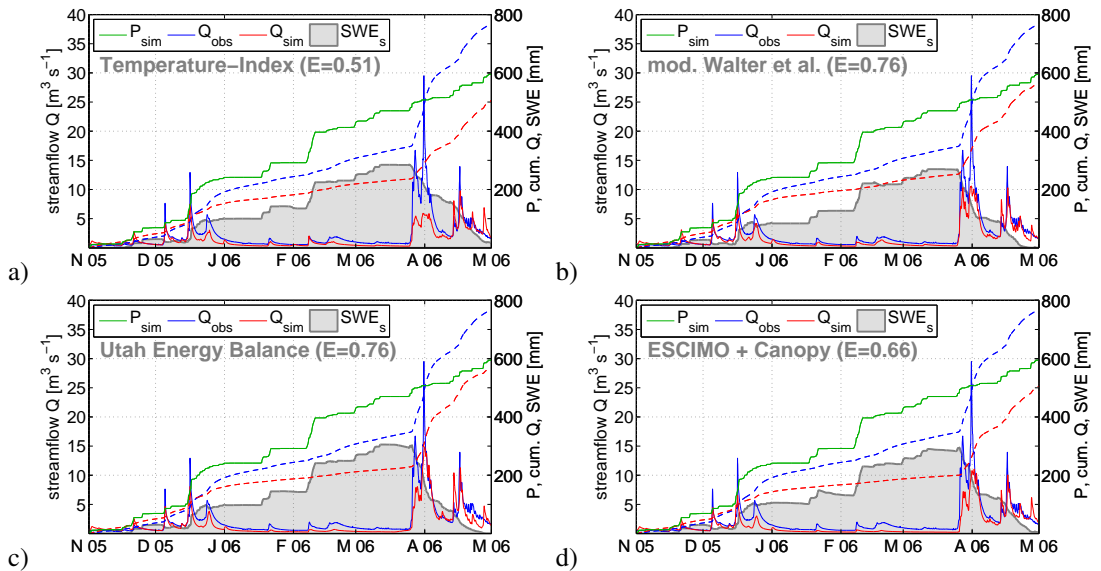


Figure 12. Streamflow simulations at the catchment scale using downscaled meteorological time series for the winter season 2005/06: (a) Temperature-Index approach, (b) modified Walter et al. (2005) approach, (c) Utah Energy Balance model, and (d) ESCIMO + Canopy. For every simulation run downscaled precipitation depth (P_{sim}), observed streamflow (Q_{obs} , blue line) including observed runoff depth (blue dashed line), simulated streamflow (Q_s , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_s) are plotted. Model efficiencies E are also given for each run.

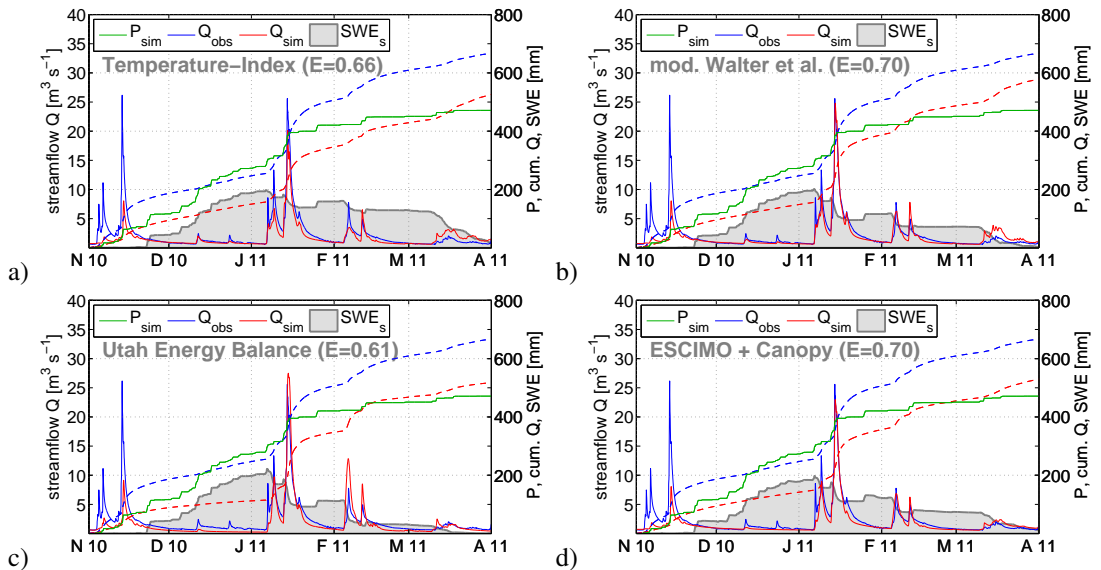


Figure 13. Streamflow simulations at the catchment scale using downscaled meteorological time series for the winter season 2010/11: (a) Temperature-Index approach, (b) modified Walter et al. (2005) approach, (c) Utah Energy Balance model, and (d) ESCIMO + Canopy. For every simulation run downscaled precipitation depth (P_{sim}), observed streamflow (Q_{obs} , blue line) including observed runoff depth (blue dashed line), simulated streamflow (Q_s , red line) including simulated runoff depth (red dashed line), and simulated snow water equivalent (SWE_s) are plotted. Model efficiencies E are also given for each run.