Hydrol. Earth Syst. Sci. Discuss., 11, 2091–2148, 2014
www.hydrol-earth-syst-sci-discuss.net/11/2091/2014/
doi:10.5194/hessd-11-2091-2014

Hydrology and Earth System Sciences
Open Access
Discussions

The Wageningen Lowland Runoff Simulator (WALRUS): application to the Hupsel Brook catchment and Cabauw polder

C. C. Brauer, P. J. J. F. Torfs, A. J. Teuling, and R. Uijlenhoet

Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, the Netherlands

Received: 2 January 2014 – Accepted: 13 January 2014 – Published: 12 February 2014

Correspondence to: C. C. Brauer (claudia.brauer@wur.nl)

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

Abstract

The Wageningen Lowland Runoff Simulator (WALRUS) is a new parametric (conceptual) rainfall-runoff model which accounts explicitly for processes that are important in lowland areas, such as groundwater-unsaturated zone coupling, wetness-dependent flowroutes, groundwater–surface water feedbacks, and seepage and surface water supply (see companion paper by Brauer et al., 2014). Lowland catchments can be divided into slightly sloping, freely draining catchments and flat polders with controlled water levels. Here, we apply WALRUS to two contrasting Dutch catchments: the Hupsel Brook catchment and Cabauw polder. In both catchments, WALRUS performs well: Nash–Sutcliffe efficiencies obtained after calibration on one year of discharge observations are 0.87 for the Hupsel Brook catchment and 0.83 for the Cabauw polder, with values of 0.74 and 0.76 for validation. The model also performs well during floods and droughts and can forecast the effect of control operations. Through the dynamic division between quick and slow flowroutes controlled by a wetness index, temporal and spatial variability in groundwater depths can be accounted for, which results in adequate simulation of discharge peaks as well as low flows. The performance of WALRUS is most sensitive to the parameter controlling the wetness index and the groundwater reservoir constant, and to a lesser extent to the quickflow reservoir constant. The effects of these three parameters can be identified in the discharge time series, which indicates that the model is not overparameterised (parsimonious). Forcing uncertainty was found to have a larger effect on modelled discharge than parameter uncertainty and uncertainty in initial conditions.

1 Introduction

Lowlands exist all over the world (often in river deltas, Fan et al., 2013). They are generally densely populated and centers of agricultural production, economic activity and transportation. Therefore, socio-economic consequences of natural hazards are

specially large in these areas. In addition, the lack of topography increases their vulnerability to flooding, climate change, and deterioration of water quality.

To mitigate natural and human disasters, hydrological models can be used by water managers as a tool for risk assessment and infrastructure design. There is growing awareness that for simulation and prediction of water and energy fluxes in lowland areas, models need to explicitly account for the dynamic groundwater table (Alley et al., 2002; Maxwell and Miller, 2005; Kollet and Maxwell, 2006; Bierkens and van den Hurk, 2007; Maxwell and Kollet, 2008). In many modelling approaches, existing models of vertical water movement in the unsaturated zone are coupled to groundwater models which simulate the horizontal flow (e.g. Gilfedder et al., 2012; Zampieri et al., 2012). This approach, however, has clear limitations in flat lowland areas where the shallow groundwater table (< 2 m below surface) often rises to within the unsaturated model domain, or even to the land surface (Appels et al., 2011; Brauer et al., 2011). In addition, surface water networks are generally dense and surface water levels influence drainage fluxes and groundwater levels. These groundwater–surface water interactions are important in both freely draining catchments and polders. Not surprisingly, hydrological models with a more traditional structure (i.e. without coupling and feedbacks) often fail to reproduce discharge dynamics in lowland catchments (Bormann and Elfert, 2010; Koch et al., 2013). Thus, instead of coupling existing models, hydrological models for application in lowland areas should be derived from a conceptually sound and strong coupling between groundwater and the unsaturated zone as well as between groundwater and surface water.

We developed a rainfall-runoff model for application in lowland areas. This model, the Wageningen Lowland Runoff Simulator (WALRUS) is described in detail in the accompanying paper (Brauer et al., 2014). The structure of WALRUS (see Fig. 1) is different from that of traditional lumped rainfall-runoff models. Firstly, the unsaturated and saturated zones are tightly coupled, so that any increase in groundwater level automatically leads to a decrease in unsaturated zone thickness and vice versa. Secondly, the model conceptualises the varying contribution of fast flowroutes through a wetness-dependent

divider (inspired by Stricker and Warmerdam, 1982). Finally, the model explicitly accounts for groundwater–surface water interaction through the inclusion of a surface water reservoir, which represents the channel network. This allows for negative feedbacks on subsurface flow during peak discharges or as a result of surface water supply.

WALRUS consists of three reservoirs: a soil reservoir (including vadose zone and groundwater), a quickflow reservoir and a surface water reservoir (Fig. 1). At the land surface, water is added to the different reservoirs by precipitation (P). A fixed fraction is led to the surface water reservoir (P_S). The soil wetness index W determines which fraction of the remaining precipitation percolates slowly through the soil matrix (P_V) and which fraction flows towards the surface water via quick flow routes (P_Q). Water is removed by evapotranspiration from the vadose zone (ET_V) and surface water reservoir (ET_S). The vadose zone is the upper part of the soil reservoir and extends from the soil surface to the dynamic groundwater table (d_G), including the capillary fringe. The dryness of the vadose zone is characterised by a single state: the storage deficit (d_V), which represents the effective thickness of empty pores (or the amount of water necessary to saturate the profile). It controls the evapotranspiration reduction (β) and the wetness index (W). The phreatic groundwater extends from the groundwater depth (d_G) downwards, thereby assuming that there is no shallow impermeable soil layer and allowing groundwater to drop below the bottom of the drainage channels (c_D) in dry periods. The groundwater table responds to changes in the storage deficit and determines, together with the surface water level, groundwater drainage or infiltration of surface water (f_{GS}). All water that does not flow through the soil matrix, passes through the quickflow reservoir to the surface water (f_{QS}). This represents macropore flow through drainpipes, animal burrows and soil cracks, but also local ponding and overland flow. The surface water reservoir has a lower boundary (the channel bottom c_D), but no upper boundary. Discharge (Q) is computed from the surface water level (h_S). Water can be added to or removed from the soil reservoir by seepage (f_{XG}) and to/from the surface water reservoir by surface water supply or pumping (f_{XS}). Model equations and abbreviations of variables used in this paper are listed in Table A1. For a more detailed

model description, see the accompanying manuscript available at Geoscientific Model Development Discussions (Brauer et al., 2014).

Whenever models are developed from a certain conceptualization of reality, they should be thoroughly tested under different circumstances to find out whether the model yields the intended outcome and to thoroughly understand the feedbacks between states, fluxes and parameters (e.g. Klemeš, 1986; Oreskes et al., 1994; Refsgaard and Knudsen, 1996; Beven, 2007; Kavetski and Fenicia, 2011). Evaluation of rainfall-runoff models can focus on (1) performance, often measured with objective functions such as the Nash–Sutcliffe efficiency, (2) uncertainty in parameter values and model structure or (3) realism of the simulated processes, by comparing to the modeller's understanding of the hydrological system and the intended function of model components (Wagener, 2003). Here we focus our evaluation on these different aspects.

In this paper we will use hydrological measurements from one freely draining catchment and one polder with controlled water levels (Sect. 2) to evaluate the performance of WALRUS during calibration (Sect. 3) and several validation runs (Sect. 4). In Sect. 5, we examine sensitivity of WALRUS to parameters, objective functions used for calibration and default functions. In Sect. 6 we investigate the effect of uncertainty in forcing, initial conditions and parameters on modelled discharge.

2 Field sites

Lowland catchments can be divided into freely draining areas with (mildly) sloping land surfaces and groundwater tables and (nearly) flat areas (called polders) where water levels are controlled by pumping water from and supplying water to a man-made drainage network. In reality, the distinction between freely draining and controlled areas is less clear, since water levels in freely draining lowland catchments are often controlled as well, e.g. by man-made channels, drainpipes, (adjustable) weirs and redirection of surface water during wet periods.

In the Netherlands, a distinction can be made between the freely draining High Netherlands (above mean sea level, although this can still be considered lowland) in the east and south of the country and the Low Netherlands with controlled water levels (below, or a few meters above, mean sea level) in the west and north (Fig. 2). Some areas with deep groundwater tables (> 10 m) exist in the far south (Limburg) and the old glacier ridges in the middle of the country (e.g. Veluwe).

Since WALRUS should be applicable to both the High Netherlands (except for the far south and the glacier ridges) and the Low Netherlands, we test it for two field sites that are characteristic for these areas and for which observations of precipitation, evapotranspiration, discharge, groundwater and soil moisture are available. The Hupsel Brook catchment is located in the relatively high eastern part of the Netherlands and the Cabauw polder in the low-lying western part (Fig. 2). Both sites have an extensive drainage network with ditches and drainpipes and have very limited or no relief, which makes it impossible to use classical topography-driven models (e.g. Beven and Freer, 2001a).

2.1 Hupsel Brook catchment

The Hupsel Brook catchment has been a well-known field site for hydrological studies since the 1960s. It has been used for studies on evapotranspiration (Stricker and Brutsaert, 1978), soil physical properties (Hopmans and van Immerzeel, 1988; Hopmans and Stricker, 1989), rainfall-runoff modelling (Stricker and Warmerdam, 1982; Bierkens and Puente, 1990) and relations between flow routes and water quality (Van den Eertwegh, 2002; Rozemeijer et al., 2010; Van der Velde et al., 2012). The catchment of 6.5 km^2 is slightly sloping (0.8 %). Its soil consists of a loamy sand layer (with some clay, peat and gravel) of 0.2 to 10 m thickness on an impermeable clay layer of more than 20 m thickness (Table 1). A more detailed catchment description can be found in e.g. Brauer et al. (2011) and Van der Velde et al. (2010).

No surface water is supplied upstream in the Hupsel Brook catchment and the elevations of several weirs and flumes in the catchment are fixed. Some small water courses

(large gullies) cross the catchment boundary (Fig. 2), but these only carry water in winter. The catchment boundary is based on a steady state groundwater map simulated with MODFLOW (Van der Velde et al., 2012), but in reality the boundary is believed to shift slightly during the year, depending on the catchment wetness and slopes of the active flow paths (groundwater gradient, drainpipe slope or channel slope). There may be some lateral groundwater flow across the catchment boundary, but this is assumed to be small in comparison with the other water balance terms.

In the Hupsel Brook catchment many hydrological variables have been measured since the 1960s, in different periods and with varying frequencies. Daily data of precipitation (P), potential evapotranspiration (ET_{pot}) and discharge (Q) are available since 1976 and hourly data since April 1979, with a gap between March 1987 and February 1992. For 8 % of the hours in the periods 1979–1987 and 1992–2013, at least one of the variables P, ET_{pot} or Q was missing.

Precipitation was measured with a rain gauge located at the meteorological station in the catchment (Fig. 2). Daily values of potential evapotranspiration (ET_{pot}) have been computed using data from the same meteorological station. Before 1988 the method of Thom and Oliver (1977) has been used and since 1989 the method of Makkink (1957). For our approach daily sums of ET_{pot} have been disaggregated to hourly sums by multiplication with the relative contribution of hourly global radiation sums to the daily global radiation sums. During the growing seasons (15 April–14 September) of 1979 through 1982 daily sums of actual evapotranspiration (ET_{act}) have been computed with the energy budget method: net radiation was measured and wind and temperature profile observations were used to estimate the sensible and ground heat fluxes. Evapotranspiration was then estimated as residual of the energy budget (for more information see Stricker and Brutsaert, 1978).

Discharge was measured by a type of H-flume at the catchment outlet (Brauer et al., 2011). Groundwater data were collected continuously at the meteorological station between 1976 and 2006. In addition, groundwater and soil moisture were measured intermittently at additional locations. From 1976 through 1984, soil moisture content and

groundwater level were measured biweekly at 6 sites (circles in Fig. 2). Soil moisture content was measured with a neutron probe at 12 depths, ranging from 0.15 to 2.05 m. Since January 2012, groundwater levels were measured hourly at 4 locations (diamonds in Fig. 2). Additional groundwater and soil moisture data are available from a field next to the meteorological station for a period around an extreme rainfall and flood event in 2010.

2.2 Cabauw polder

The Cabauw polder area considered as a catchment in this study is 0.5 km^2 and part of a larger polder (Table 1). Its soil consists of heavy clay on peat and is characterized by an intensive drainage network of channels and drainpipes. Water is supplied upstream into the area from a more elevated water course through a variable inlet controlled by the water authority and through two small pipes with relatively constant discharge (Fig. 2). Surface water supply is necessary to raise groundwater levels for optimal crop growth and to prevent peat oxidization, while maintaining surface water flow velocity to avoid algal blooms in standing water. Downstream of the outlet is a larger water course, from which water is pumped into the river Lek (a large branch of the Rhine delta). It is important to note that there is no pumping station within the catchment and hence drainage is driven by gravity. The surface water levels are regulated by two weirs, which are set 10 cm higher in summer than in winter. The variable inlet is used to maintain these surface water levels. Surface water levels vary to keep groundwater at an optimal depth: deep enough to avoid waterlogging and to provide a firm ground for tractors (wet clay and peat are too unstable) and high enough to avoid oxidization of peat and plant water stress. In winter, groundwater levels are convex between ditches because precipitation exceeds evapotranspiration and as a consequence groundwater flows towards the ditches. In summer, groundwater levels are concave between ditches because evapotranspiration exceeds precipitation and hence water infiltrates from ditches into the soil.

The "catchment" is part of the Cabauw Experimental Site for Atmospheric Research (CESAR), which is well-known in the international meteorological community (Russchenberg et al., 2005; Van Ulden and Wieringa, 1996; Chen et al., 1997; Leijnse et al., 2010). The site is maintained by the Royal Netherlands Meteorological Institute (KNMI) and a consortium of 8 Dutch institutes (including Wageningen University). The site contains a 213 m high measurement tower, a separate flux tower for studies on land surface–atmosphere interaction (a FLUXNET location, Baldocchi et al., 2001), and many additional instruments. Extensive summaries can be found in Russchenberg et al. (2005) and Leijnse et al. (2010). Data from Cabauw have been used in hydro(meteoro)logical studies to estimate land-surface fluxes with SWAP (Gusev and Nasonova, 1998), to investigate the effect of spatial variability in rainfall on soil moisture, groundwater and discharge with SIMGRO (Schuurmans and Bierkens, 2007), and to assess the transferability of land-surface hydrology models (Devonec and Barros, 2002).

Precipitation is measured with an automatic rain gauge, potential evapotranspiration is estimated with the approach of Makkink (1957) and actual evapotranspiration is determined by measuring net radiation, ground heat flux and Bowen ratio (with an eddy covariance set-up) and closing the energy balance (Beljaars and Bosveld, 1997; Foken, 2008). ET_{act} estimated with this method was on average 4 % higher than ET_{pot} during well-watered conditions (meaning that the storage deficit was below 100 mm). Overestimation of the daily evapotranspiration sum may be caused by an underestimation of dew formation at night (De Roode et al., 2010). As a quick fix, we divided ET_{act} by 1.04. Using ET_{act} estimated from the eddy covariance set-up directly was not an option due to the underestimation of eddy-covariance measurements which is often reported in the literature (e.g. Twine et al., 2000) and amounts to 18 % in the Cabauw polder.

Discharge is measured since May 2007 using a V-notch weir (downstream of the variable inlet) and trapezoidal Rossby weir (outlet), of which the stage–discharge relations have been obtained by laboratory calibration. The uncertainty associated with the

discharge measurement of surface water supply is large because the V-notch weir was often submerged due to the small topographical gradient. In addition, the two small inlets (pipes) with relatively constant discharge were maintained by local residents and could not be measured continuously.

Groundwater levels have been measured since August 2003 with nine piezometers in the transect in the southeast corner of the catchment (Fig. 2): 5 automatically (4 h resolution) and 4 manually (biweekly resolution). Soil moisture contents have been measured daily between November 2003 and August 2010 with a TDR set-up developed by Heimovaara (1990), consisting of four arrays of six sensors between 5 and 73 cm below the soil surface.

There is likely groundwater flow into the catchment from the nearby river Lek (1 km to the south), of which the water level is variable and on average about 2 m higher than the water levels in the catchment (and 0.2–1.5 m above mean sea level). The top soil consists of a mixture of clay and peat and is not permeable enough for significant groundwater flow, but locally flow may occur through buried river sands (National Institute for Drinking Water Supply, 1982). Because no seepage data are available, we estimated the seepage as residual of the water budget of the year November 2007–October 2008 (also used for calibration, see Sect. 3) and assuming a constant seepage flux year-round. This seepage estimate amounts to about 7 % of the annual water budget.

2.3 Climatology

Since the Hupsel Brook catchment and Cabauw polder area are located only 120 km apart, the climate is quite similar: annual precipitation is around 800 mm and the annual potential evapotranspiration amounts to 600 mm (Table 1). The actual evapotranspiration (ET_{act}) in the Hupsel Brook catchment is usually within 5 % of ET_{pot} (based on 4 yr of combined observations). In the Cabauw polder, shallow groundwater tables prevent a strong soil moisture limitation on evapotranspiration (Brauer et al., 2014). Precipitation occurs on 50 % of the days, but quantities are typically low: on 15 % of the days

more than 5 mm was measured and on 5 % more than 10 mm. During 11 % of the hours precipitation was observed, of which 75 % with accumulations less than 1 mm. Hourly rainfall sums above 10 mm occur on average 3 times per year (at a given location and based on clock hours rather than a moving window).

Snow is of limited importance, even though freezing conditions are common. Sub-zero daily average temperatures occur on average on 28 (Hupsel) and 18 (Cabauw) days per year, leading to freezing of ponds on the land surface, water in soil cracks and drainpipes, and the top layer of slowly flowing or standing surface water. On the majority of these days, the daily maximum temperature is above zero, leading to daily cycles of freezing and thawing. Cold winter conditions are often caused by persistent high pressure systems with little precipitation: on average 0.4 (Hupsel) and 0.3 (Cabauw) mm of precipitation on days with daily mean temperature below zero, leading to on average 12 (Hupsel – 1.5 % of total P) and 6 (Cabauw – 0.8 % of total P) mm of precipitation annually.

It should be noted that water input from dew can be considerable. Jacobs et al. (2006, 2010) estimated dew to amount to 4.5 % of the annual precipitation sum at Wageningen (located roughly halfway between the Hupsel Brook catchment and the Cabauw polder). Unfortunately, dew measurements were not available for either catchment. Therefore, dew is not considered separately in the water balance, but assumed to be included in the rain gauge measurements.

Water balance terms show seasonal variation (Fig. 3). Evapotranspiration exceeds precipitation between April and August, which means that excess water stored in winter and, in the case of the Cabauw polder, surface water supply f_{XS} and seepage f_{XG} are used in summer for both Q and ET. The influence of water management in the Cabauw polder is clearly visible: discharges remain high in summer due to surface water supply, on 6 May 2008 discharge suddenly dropped to zero as a result of the increase of weir elevation, and on 16 November 2007 and 15 October 2008, discharge increased because the weir was lowered. The surface water supply flux in the Cabauw polder is large and variable and can reach 800 mm in some years. In the Cabauw polder, there

is always water in all branches of the surface water network, whereas the headwaters of the Hupsel Brook frequently run dry. Discharge at the outlet of the Hupsel Brook catchment dropped to zero during three months in 1976, a month in 1982 and several shorter periods in 1983, 1988 and 2011.

3 Calibration

Because geology, slope and drainage densities differ between catchments, the model parameters expressing the effects of these characteristics differ as well. Although the parameters have physical connotations, they are effective values representing the entire catchment (including the effect of heterogeneity). Therefore, they cannot be estimated from field measurements directly, but have to be calibrated. Fitting simulations to observations yields catchment-specific parameter values, which (as they are assumed to be time-invariant) can be used to simulate discharge during other periods.

For both the Hupsel Brook catchment and the Cabauw polder, we optimized four model parameters: the wetness index parameter c_W, vadose zone relaxation time c_V, groundwater reservoir constant c_G and the quickflow reservoir constant c_Q (see Fig. 1 and Table A1 for a complete overview of all model variables, parameters and relations). We used the stage–discharge relations of the outlet weirs (which were calibrated in the laboratory) and channel depths c_D of 1500 mm (estimated from observations). The weir level $h_{S,\,min}$ in the Cabauw polder was set to 500 (winter) and 600 (summer) mm from the channel bottom (based on field estimates). We used default functions for $W(d_V)$, $d_{V,\,eq}(d_G)$ and $\beta(d_V)$ and soil physical parameters b, ψ_{ae} and θ_s based on observations in the Hupsel Brook catchment and the Cabauw polder (see Brauer et al., 2014).

For the calibration, we used hourly data of the periods November 2011–October 2012 (Hupsel) and October 2007–September 2008 (Cabauw). Unfortunately, it was not possible to use the same period for both catchments, since time series were not continuous. Both periods are not exceptionally dry or wet and do not contain long periods of sub-zero temperatures (except February 2012). Several choices of objective

functions are compared in Sect. 5.5, but a classical Nash–Sutcliffe efficiency of the discharge was used as main objective function (Nash and Sutcliffe, 1970).

We used a particle swarm optimization technique by Zambrano-Bigarini and Rojas (2013), called HydroPSO, which is less sensitive to discontinuities in the response surface (i.e. due to thresholds in the model) and more likely to find a global optimum than other gradient-based methods. The parameter values obtained with this HydroPSO calibration are used throughout the paper. A comparison of optimization algorithms is outside the scope of this paper. In addition to the calibration with HydroPSO, a Monte Carlo analysis was used to explore uncertainty in and dependency between parameters (in Fig. 4, and Sects. 5.4 and 6.1). For the Monte Carlo analysis, we generated 10 000 random parameter sets with ranges 100–500 mm (c_W), 0.1–20 h (c_V), 0.1–150 mm h (c_G) and 1–100 h (c_Q).

3.1 Calibrated parameter values

The optimal parameter values found with HydroPSO and the relations between parameter values and Nash–Sutcliffe efficiencies obtained with the Monte Carlo analysis are shown in Fig. 4. Finding optimal parameter values is not trivial (e.g. Beven and Freer, 2001b; Melsen et al., 2013). We used HydroPSO to obtain one optimal parameters set, but the dotty plots show that equally good results (in terms of Nash–Sutcliffe efficiency) could have been obtained with different combinations.

When comparing the Cabauw polder to the Hupsel Brook catchment, differences in parameter values can be observed and explained. Parameters c_V, c_G and c_Q are higher, indicating that all flow is slower. Parameter c_W is smaller, causing earlier activation of quick flowroutes (at higher storage deficits). Compared to the Hupsel Brook catchment, the clayey soil in the Cabauw polder is less permeable, leading to slower groundwater flow (c_G) and a slower response of groundwater to changes in the unsaturated zone (c_V). There are more cracks, gullies and drainpipes per unit area (c_W), but quickflow is slower because slopes of land surface (overland flow) and drainpipes are lower (c_Q). It is not a coincidence that the drainage density increases when

permeability decreases. Farmers install drainpipes and dig gullies when ponding hampers agricultural activities, animals (moles, mice and muskrats) dig more burrows to drain their dens and cracks occur more quickly in clayey soils.

3.2 Calibrated results

Discharge is reproduced well during the calibration period, both for peaks in winter and for low flows and small peaks in summer (Fig. 5). Nash–Sutcliffe efficiencies of 0.87 for the Hupsel Brook catchment and 0.83 for the Cabauw polder are reached. This shows that the model with the optimal parameters is able to capture the hydrological response of lowland catchments.

In February 2011 the headwaters of the Hupsel Brook were frozen, which caused a decrease in observed discharge. Because WALRUS in the present form does not take freezing conditions and snow into account, the simulated discharge does not decrease as quickly.

WALRUS simulates the discharge in summer relatively well. Although the groundwater level dropped below the channel bottom (in agreement with reality), the channel did not run dry, because both discharge and infiltration of surface water decrease rapidly at low water levels. Only evapotranspiration can empty the channel completely. During summer field visits, we frequently observed that, while a large part of the surface water network is dry, some storm water is still discharged at the outlet after rain events. Even when the soil is dry, some quickflow will occur close to the ditches or over paved areas.

The modelled groundwater depth d_G shows seasonal variation, but does not respond quickly to rainfall events. In the model, percolating water is significantly delayed in the vadose zone and the dynamic response to rainfall is modelled by the quickflow reservoir. The groundwater depth does influence the catchment's quick response to rainfall events, because when groundwater is shallow, percolation is slow and storage deficits are small, resulting in a high wetness index and a large portion of the rain being led through the quickflow reservoir.

Groundwater drainage (f_{GS}) shows both long-term and short-term dynamics. The seasonality in f_{GS} is caused by seasonality in groundwater levels, which are higher in winter due to the precipitation surplus. The quick decreases are caused by fluctuations in surface water level, which rise and fall rapidly after rainfall events. This shows that the groundwater–surface water feedback is implemented appropriately: during discharge peaks, drainage is limited by high surface water levels.

The surface water levels are much more constant in the Cabauw polder than in the Hupsel Brook catchment. Surface water supply prevents headwaters from running dry in summer. In addition, the Cabauw polder has a five time larger fraction of surface water, which acts as a buffer and absorbs inflow peaks caused by rainfall events.

3.3 Water budget

In the Hupsel Brook catchment, quickflow (f_{QS}) accounts for 70 % of total drainage ($f_{GS} + f_{QS}$; Table 2). This is consistent with the important role of quickflow found in previous studies. Van der Velde et al. (2011) measured drainpipe flow in one field in the Hupsel Brook catchment and found that the contribution of drainpipe flow (one of the components of quickflow) to the total drainage was 80 % for that field and estimated it to be 25–50 % for the entire catchment.

In the Cabauw polder, the contribution of groundwater drainage is limited. However, the groundwater depth plays an important role in dividing the water between the soil reservoir and the quickflow reservoir.

In both catchments the change in storage in the soil reservoir is considerable: −54 mm in the Hupsel Brook catchment and −110 mm in the Cabauw polder. Observations of discharge in the Hupsel Brook catchment (as a proxy for storage) and soil moisture in the Cabauw polder show that both years chosen for calibration ended drier than they started. However, the decrease in storage in the Cabauw polder was overestimated.

Evapotranspiration reduction is negligible in the Cabauw polder (604/607 = 0.5 %), but significant in the Hupsel Brook catchment (531/587 = 10 %), caused by the larger soil moisture deficits.

4 Validation

The parameter values obtained during the calibration runs described in the previous section were used in validation studies for whole years, a short period to focus on groundwater dynamics, major flood and drought events, and a case with management operations. We altered the initial groundwater depth for every validation run to match the real catchment wetness at the start of each validation period (because in contrast to parameters, initial conditions are not time-invariant).

4.1 Validation on yearly timescale

For both catchments, we selected one year for which actual evapotranspiration, soil moisture and groundwater data were available: the period 15 April 1979–15 April 1980 for the Hupsel Brook catchment and 1 November 2008–1 November 2009 for the Cabauw polder. These additional data are used to test whether the model only reproduces the observed discharge or also the hydrological processes involved. The requirement of these additional data and allowing no gaps limited the choice of years for validation studies to one or two (different) years for each catchment and we chose years that are not exceptionally dry or wet and do not contain long periods of freezing conditions (except January 2009).

Model results and measurements are shown in Fig. 6 and some annual sums of water balance terms are shown in Table 2. For both catchments, Nash–Sutcliffe efficiencies are lower for the validation runs than for the calibration runs, but still acceptable: they decrease from 0.87 to 0.74 for the Hupsel Brook catchment and from 0.83 to 0.76 for the Cabauw polder. This relatively small decrease in performance indicates

that the model is parsimonious. In both catchments the highest discharge peaks are underestimated.

During a field visit in the Cabauw polder in December 2008, a culvert was found clogged with loose vegetation, reducing discharge capacity and raising water levels upstream. When the blockage was removed, the water stored upstream was released, leading to a sharp discharge peak.

In the Hupsel Brook catchment, modelled storage deficits and groundwater depths fall within the range of observations of the different stations, but in the Cabauw polder, the groundwater depth is overestimated at the end of the year. Including groundwater levels in the calibration procedure (multi-objective calibration) may lead to better estimates.

4.2 Groundwater dynamics

To evaluate the modelled response of groundwater to rainfall events, we selected two periods for which groundwater data with high temporal resolution were available (for locations, see Fig. 2). In the Hupsel Brook catchment, the piezometers were well distributed over the catchment, leading to a large variability in observed groundwater depths and dynamics (Fig. 7). The piezometer with the shallowest and most dynamic groundwater table is located close to the surface water network, in the part of the catchment where the aquifer is thin. As a consequence, ponding and overland flow occur relatively quickly and the drainage density is large. The two piezometers with the deepest and least dynamic groundwater tables are located where the aquifer is thick and permeable. The relatively thick unsaturated zone attenuates infiltrating water and ponding and overland flow hardly ever occur.

The modelled groundwater depth falls within the range of observations in the Hupsel Brook catchment, but hardly varies in time. This is caused by the function of the groundwater reservoir in WALRUS. The groundwater reservoir only simulates the seasonal groundwater dynamics for groundwater drainage, while the quickflow reservoir accounts for the dynamic response to individual rainfall events. Observed groundwater

levels measured at different locations are the results of different contributions of quickflow and groundwater flow. The most dynamic groundwater tables in Fig. 7 can be compared to a combination of the quickflow and groundwater reservoirs, while the least dynamic groundwater tables mainly represent the groundwater reservoir alone. In the Cabauw polder, observed groundwater depths are much more variable than simulated ones (Fig. 7). This indicates that the contribution of quickflow is large in all measured locations.

By using two reservoirs to simulate the characteristic “coupled dynamics” of groundwater tables, the discharge can be reproduced well. Using groundwater time series for calibration or validation is possible, but not trivial (as is the case for all lumped rainfall-runoff models). In addition, piezometers are often situated in locations with large seasonal dynamics. The 6 piezometers used in the Hupsel Brook catchment in the 1970s and 1980s overestimate the variation in total catchment storage (reflecting the seasonal variation), which may be caused by the installation of piezometers in the centre of fields rather than near the channels (Brauer et al., 2013a).

4.3 Extreme rainfall and flash flood

On 26 August 2010, an extreme rainfall event occurred in the Hupsel Brook catchment (Brauer et al., 2011). In 24 h, about 160 mm rainfall was observed, corresponding to a return period of more than 1000 years. This resulted in soil saturation, overland flow and inundation. Some of the lessons learnt from the analysis of this flash flood event, such as the importance of the groundwater–surface water feedback and wetness-dependent flowroutes, were taken into account during the development of WALRUS (Brauer et al., 2014). The flood in August 2010 has triggered a model intercomparison study (without WALRUS) initiated by the Dutch Hydrological Association (NHV), in which teams from Dutch consultancy firms, institutes and universities participated (NHV, 2013).

We used the calibrated model to simulate this extreme event (not part of the calibration period). We used the mean of the observed groundwater depths as initial

groundwater level for the simulation, mimicking a real flood forecasting situation, where information about the discharge in the future is not available. This yielded a relatively good discharge response, considering that neither parameters nor initial conditions were fitted to the data of the event (Fig. 8). The Nash–Sutcliffe efficiency, however, was relatively low (0.64), mostly because the timing of the peak was slightly off. The total discharge was overestimated by WALRUS: 98 mm, compared to 87 mm observed (runoff ratios of 0.50 and 0.44). The peak discharge was simulated accurately, but the top was flatter than observed, because surface water levels exceeded the soil surface and water flowed overland to the groundwater reservoir. The peak was capped at the bankfull discharge, which has been defined by the stage–discharge relation. In reality, the peak has been capped in many channels at different heights and the combination of many thresholds probably led to the observed smooth curve. Altogether, WALRUS performs better than many models used in the intercomparison study, in which peak discharges were reported ranging from 0.2 to 10 mm h^{-1} (NHV, 2013).

In the model, complete catchment saturation is never reached in the soil reservoir. However, the wetness index reached 0.87, which means that 87 % of the precipitation is led through the quickflow reservoir. So, even though the effective groundwater depth did not reach the soil surface, ponding and overland flow occurred in a large part of the catchment. The wetness index reproduces an increasing fraction of ponding and overland flow in local depressions in the landscape and near the channels, while local elevations in the landscape remain unsaturated. With this technique, this lumped model can account for spatial variability of groundwater depths.

Observations show that groundwater reached the soil surface before overland flow occurred (Fig. 8), while according to the model water flowed overland from the surface water to the soil reservoir. Of course, the observations are point measurements and it is likely that areas closer to the channels have flooded before reaching saturation, while local elevations in the landscape remained dry. This example shows that relating modelled (catchment effective) variables to point measurements of groundwater depth and soil moisture content is not trivial.

The satisfactory results indicate that WALRUS can be applied for flood forecasting. The initial storage deficit (and groundwater level) has a large influence on the simulated discharge. It determines when quickflow starts, when the surface water reaches the soil surface and when overland flow towards the groundwater reservoir starts. State updating could reduce the predictive uncertainty resulting from uncertainty in initial conditions when used in an operational flood forecasting/early warning system.

4.4 Extreme summer drought

In 1976, much of western Europe including the Netherlands experienced one of the worst summer droughts in recent history (Van Huijgevoort et al., 2013). The annual precipitation sum in the Hupsel Brook catchment was 549 mm for the whole year, compared to 790 mm on average. High evapotranspiration accelerated the development of large storage deficits (Teuling et al., 2013). During this summer, intensive field observations in the Hupsel Brook catchment took place (Stricker and Brutsaert, 1978). Because hourly data were not available, daily data were used as input.

In general, the discharge was simulated well. The simulated initial recession in April and May is too steep and the response to rainfall events in late May and June is too strong, but the limited response to rainfall later on is simulated correctly. It should be noted that extreme drought conditions can temporarily change soil properties and hydrologic response (Seneviratne et al., 2012), which might explain the slight model mismatch in this period. Even during this extremely dry summer, some discharge was observed after rainfall events. This is simulated well by WALRUS, where a small portion of the rainfall is led through the quickflow reservoir, mimicking runoff from paved surfaces or through depressions near the surface water network. This shows the added value of the quickflow reservoir and the surface water reservoir – in models with only a groundwater reservoir, all rainfall would infiltrate into the soil and discharge would remain zero.

No observations of actual evapotranspiration were made before 15 April and after 15 September. It was assumed that it would not deviate much from its potential value

in winter, which is confirmed by the data: ET_{act} is close to ET_{pot} in late April and early September (Fig. 9). Between May and August, a large evapotranspiration reduction is observed. The model simulates the evapotranspiration reduction well, apart from a slight underestimation at the start and a slight overestimation at the end of the period. For the whole period shown in Fig. 9, modelled evapotranspiration reduction was 30 % compared to 26 % observed. Depletion of groundwater and soil moisture were slightly underestimated, but fall well within the range of observed values.

4.5 Effect of water management

WALRUS can incorporate water management operations, which is important if it is to be used in human-influenced lowlands. To investigate if the model can also be used for water management scenario analyses or to separate natural and human effects on the hydrological system (see Van Loon and van Lanen, 2013, and references therein), we simulated the change in the hydrological variables as a result of water management operations in the Cabauw polder. As mentioned in Sect. 2.2, surface water levels in the Cabauw polder are controlled by adjusting weir elevations and regulating surface water supply.

In Fig. 10 model results are shown for situations with and without management operations, being lowering of the weir and increasing surface water supply. Observations of the actual situation (i.e. with management operations) are shown as well. The model reproduces the discharge response to water management operations well, although time delays become visible when we zoom into the short time windows in Fig. 10. Lowering the weir causes a discharge peak, as the water stored in the top 10 cm of the surface water is released quickly (left panels). As 5 % of the catchment is covered by surface water, this amounts to ($0.05 \times 100 =$) 5 mm of water averaged over the catchment area. A sudden increase in surface water supply also leads to a discharge rise, but less rapidly, because first the extra water has to be distributed over the surface water network. This delay represents mostly the response time of the surface water system, which is determined by the surface water area fraction and the stage–discharge

relation. This indicates that a surface water reservoir (such as incorporated in WALRUS) is necessary to simulate the effect of the water buffer.

5 Sensitivity analyses

We performed three types of analyses to assess the sensitivity of modelled discharge to changes in parameter values: Sect. 5.1 focusses on parameter identifiability through a time series analysis, Sect. 5.2 focusses on the parameter sensitivity with a novel statistical technique (DELSA) and Sect. 5.4 focusses on parameter uncertainty and dependence using an analysis of response surfaces. In addition, we investigate the sensitivity to the choice of objective function for calibration in Sect. 5.5 and choice of user-defined parameterizations in Sect. 5.6 (described in Brauer et al., 2014, and listed in Table A1).

5.1 Parameter identifiability

Calibration is improved and the risk of equifinality reduced when the influence of each parameter can be distinguished in the discharge time series. In this section, the derivative of discharge to each of the parameters ($\partial Q/\partial c$) is determined, keeping the others fixed. This sensitivity is approximated by a numerical difference $\left(\frac{Q(c+\Delta c)-Q(c)}{\Delta c}\right)$, with $\Delta c = 10^{-4}$.

The parameter sensitivity is plotted in Fig. 11 for all four calibration parameters, focusing on the Hupsel Brook catchment in December 2011 and January 2012 (part of Fig. 5). To facilitate comparison, we scaled each sensitivity time series with the parameter value in question.

The sensitivity series of c_W, c_G and c_Q are clearly different enough to make these parameters identifiable. Moreover, the differences can be understood. Sensitivity to the wetness index parameter c_W is large at the start of the period, when wetness index (W) is increasing after the dry summer period and decreases as the winter progresses. With

a larger value of c_W, W will be larger at the same value of d_V, leading to more quickflow and higher discharge peaks initially. Because less water is led to the soil reservoir in comparison to the original simulation, d_V decreases less quickly and the same value of W is reached with a different combination of c_W and d_V.

As c_V, c_G and c_Q cause delay and attenuation, an increase in these parameters dampens the discharge signal. The effect of c_Q is easily understood, because there are no direct feedbacks between the quickflow and surface water reservoirs: an increase in c_Q causes a lower and longer discharge peak. The peak decreases (negative $\Delta Q/\Delta c$) and the tail height increases (positive $\Delta Q/\Delta c$). A larger value of c_G causes a decrease of groundwater drainage and lower discharge between peaks. Surface water infiltration during peak flows is limited as well, leading to increased discharge peaks.

Discharge is about 1000 times less sensitive to the vadose zone relaxation time parameter c_V than to the other parameters (see the length of the vertical coloured bars in Fig. 11). The time series of sensitivity to c_V in inversely proportional to the sensitivity to c_W. This indicates that it is impossible to distinguish the effect of c_V in the discharge time series and that calibration of this parameter with discharge data alone is impossible.

5.2 Parameter sensitivity

A more sophisticated method to determine the sensitivity of WALRUS to model parameters is the Distributed Evaluation of Local Sensitivity Analysis (DELSA, see Rakovec et al., 2014, for a complete explanation of the method). In short, this hybrid local-global sensitivity method decomposes the variance of a performance measure into contributions from individual parameters using multiple evaluations of local parameter sensitivities, which are distributed throughout parameter space. The current implementation of the DELSA method provides first-order sensitivities for each of the model parameters. This means that only main effects on the total variance are captured and no parameter interactions are considered. In addition, the DELSA values conveniently scale between 0 and 1 and – when all variance is explained by one parameter, its DELSA sensitivity

is 1. Finally, the advantage of DELSA is that a rather small sample size (yielding low computational cost) provides robust results.

To compute the DELSA values, we initially created 100 parameter sets (the base set). Next, we took the base set and perturbed one of the parameters, which we repeated for each of the four parameters. We ran WALRUS with these 500 sets and we evaluated the model output using three performance measures: the sum of squares (SS) of Q, of Q^2 (to focus on peaks) and of $\sqrt{Q}$ (to focus on low flows). For each of the four parameters, the DELSA sensitivity was computed from the difference in parameter value and model performance between the base run and the run with the perturbed parameter.

Figure 12 shows boxplots of the obtained DELSA sensitivity for each parameter and each performance measure, in which the ranges indicate the variation between parameter sets. The sensitivity to c_V is again small and the sensitivity to c_Q is only large when extra focus is placed on the peaks (SS(Q^2)). These two parameters only change the discharge temporarily, while c_W and c_G have a long-lasting effect through groundwater recharge (c_W) and recession (c_G). WALRUS is sensitive to c_W for many parameter sets (high sensitivity for low quantiles), especially when focussing on low flows (SS($\sqrt{Q}$)). Because c_W determines the amount of water that is led to the soil reservoir, and consequently the starting level of recession periods, a small change in this parameter can lead to overestimation or underestimation of baseflow.

5.3 Conclusions of parameter sensitivity analyses

In conclusion, the discharge is most sensitive to parameters c_W, c_G and c_Q. These parameters are identifiable in the discharge time series. This gives confidence that the model is not overparameterized, which facilitates calibration and reduces the risk of equifinality. For c_V, however, an optimum cannot be determined with calibration on discharge data alone for the winter periods in the Hupsel Brook catchment analysed in Sect. 5.1. This does not mean that c_V is superfluous – c_V controls the delaying

influence of the unsaturated zone, which is not visible unless one zooms in on individual discharge peaks.

5.4 Parameter uncertainty

Parameter uncertainty (or the statistics thereof) can be assessed by analysing the surface of the Nash–Sutcliffe efficiency as function of the parameters near the optimum. A first step into this direction is given in Fig. 4, where the Nash–Sutcliffe efficiency is plotted as a function of each individual parameter. This Figure was obtained from the Monte Carlo analysis with 10 000 random parameter sets (see Sect. 3). Figure 4 shows that the curvature of the Nash–Sutcliffe surface near the optimum clearly differs between parameters, leading to different uncertainties. For example, the optimum of c_Q is clearly defined, while high Nash–Sutcliffe efficiencies appear over the whole range of c_V.

To analyse the simultaneous dependence of the Nash–Sutcliffe efficiency on two parameters, we made response surfaces for all parameter combinations for both catchments using the output from the same Monte Carlo analysis. As illustration, we plotted two response surfaces for the Hupsel Brook catchment in Fig. 13. The surfaces were obtained by inverse distance interpolation of the Nash–Sutcliffe efficiencies of the Monte Carlo simulations. The response surfaces are not entirely horizontal or vertical ellipses, indicating some parameter dependence. The c_W–c_Q combination leads to a slightly tilted ellipse, indicating that their optima are positively correlated. The top of the c_G–c_Q response surface is slightly horse-shoe shaped, leading to lower Nash–Sutcliffe efficiencies around c_G = 5 h and c_Q = 15 mm h. Negative values of c_G and c_Q are not physical and were therefore not chosen in the Monte Carlo analysis. It is also visible that the parameter c_Q has a different optimum in combination with c_W (30–60 h) than with c_G (5–30 h), which hampers calibration. The other parameter combinations lead to similar response surfaces, from which we can conclude that parameters in WALRUS are not independent, but do not show strong dependencies either.

For practical applications of WALRUS this rather computationally expensive Monte Carlo analysis can be replaced with a classical linearisation of the model near the optimum and an analysis of the resulting Hessian.

5.5 Sensitivity to calibration objective function

In this Section, we evaluate the effect of the choice of objective function used for calibration on the identified model parameters. We calibrated WALRUS for the Hupsel Brook catchment using 4 performance measures: the sum of squares of (1) the discharge, (2) the square of the discharge to focus on peaks, (3) the square root of the discharge to focus on low flows and (4) the groundwater level measured at the meteorological station. Because all model variables are given as model output and because the calibration does not occur within the model, calibration criteria can be changed easily. We used the longest period for which hourly groundwater data were available and no frost occurred: 1 March 2012 to 20 January 2013.

Fitting on $\sqrt{Q}$ leads to a higher value of the quickflow reservoir constant c_Q (12 h) compared to the fit on Q (4 h) and Q^2 (1 h) (Table 3). A high c_Q causes lower and broader peaks, improving the fit of the recessions (and worsening the fit of the peaks), while a low c_Q improves the fit of the peaks. Fitting on Q^2 yields a higher c_W, causing more water to be led to the quickflow reservoir. Fitting on d_G leads to a very small c_W – all water is led to the soil reservoir to mimic the dynamics of the observed groundwater depth. The observed groundwater depth, however, is represented by a combination of the soil reservoir and quickflow reservoir rather than the soil reservoir alone (Sect. 4.2). The large value of c_Q for the fit on d_G is insignificant, because no water is led to the quickflow reservoir.

5.6 Sensitivity to default parameterisations

There are four relations between model variables which can be specified by the user and for which defaults have been implemented: (1) the wetness index relation $W(d_V)$,

(2) the evapotranspiration reduction function $\beta(d_V)$, (3) the relation between equilibrium storage deficit and groundwater depth $d_{V,eq}(d_G)$, and (4) the stage–discharge relation $Q(h_S)$ (Table A1). These parameterisations are considered to be identifiable without calibration. Neverheless, they are also prone to some uncertainty. To examine how sensitive the model is to changes in these relations (i.e. the effect of choices), we ran the model with different options for these functions, with and without recalibrating for each function.

As an example, the results for three options for $d_{V,eq}(d_G)$ are shown in Fig. 14. The default option is the relation based on a power law soil moisture profile and data from the Hupsel Brook catchment (Brauer et al., 2014). We also used the relation based on a power law soil moisture profile of loamy sand (Brooks and Corey, 1964) and a linear fit through observations in the Hupsel Brook catchment. The different relations are shown in the inset of Fig. 14.

In the top panel of Fig. 14, the same values for the four model parameters (c_W, c_V, c_G and c_Q) were used. These were obtained from calibration using the default option for $d_{V,eq}(d_G)$. The initial conditions are computed automatically for each run, assuming stationary groundwater drainage (implemented as default). In the bottom panel, the model parameters were calibrated using the $d_{V,eq}(d_G)$ function in question.

This Figure illustrates that parameters obtained using one function cannot be used directly with another function. The difference between the linear and power-law based fit on the data is limited, but for the theoretical relation peaks are strongly overestimated when the original parameter set is used. However, calibration using this relation yielded similar results.

6 Uncertainty propagation

Because WALRUS is computationally efficient, it is feasible to estimate the effect of different types of uncertainty by creating ensembles of model output. In this Section we

investigate the consequences of uncertainty in parameter values, initial conditions and forcing data.

6.1 Propagation of parameter uncertainty

To examine the effect of parameter uncertainty, we created 10 000 parameter sets randomly by selecting from uniform distributions with ranges displayed in Fig. 4. We selected the 100 sets which yielded the highest Nash–Sutcliffe efficiencies for the calibration period used in Sect. 3 (November 2011–October 2012). These 100 parameters sets were used for 100 simulations of the period April 2012–May 2012.

The range between the 10th and 90th percentile is shown in Fig. 15a. Parameter uncertainty causes the largest deviations during peak flows and decreases to almost zero during recessions. The uncertainty around the large peak of 0.07 mm h^{-1} is quite large: the range between the 10th and 90th percentile ranges from 0.004 to 0.14 mm h^{-1}.

6.2 Propagation of initial condition uncertainty

Initial groundwater depth and quickflow reservoir level can be specified by providing the fraction of discharge at $t = 0$ which originates from groundwater (G_{frac}). The remainder ($1 - G_{frac}$) is used to compute the quickflow reservoir level. To investigate the effect of these initial conditions, the model calibrated in Sect. 3 is run with an initial groundwater depth based on 0 % and 100 % of discharge originating from drainage (G_{frac} of 0 and 1).

For this catchment and period, uncertainty in initial conditions has less effect on simulated discharge than uncertainty in parameter values (Fig. 15b). The range around the large peak is 0.04–0.07 mm h^{-1}. The difference between the simulations with wet and dry initial conditions decreases in time. During this period, groundwater dropped 22 mm for the wet initial condition (100 % drainage) and 16 mm for the dry initial condition (0 % drainage) and the discharge range decreased slowly as well.

6.3 Propagation of forcing uncertainty

Precipitation time series contain errors and uncertainties which can have a large influence on model performance (e.g. Beven, 2012; Pappenberger et al., 2005; Berne et al., 2005; Tetzlaff and Uhlenbrook, 2005; Hazenberg et al., 2011). For many catchments no accurate precipitation data are available and data from rain gauges outside the catchment are used. To investigate the effect of this error, we used precipitation data from the three closest operational Dutch rain gauges with hourly resolution (30–50 km from the Hupsel Brook catchment) to run WALRUS. In addition, we used the operational weather radar of the Royal Netherlands Meteorological Institute. Radar data have been adjusted with rain gauge observations (Overeem et al., 2009). The spatial resolution was 1 km^2 (Overeem et al., 2011), leading to 7 pixels for the Hupsel Brook catchment.

The effect of different rainfall inputs on modelled discharges is very large, especially after the first peak in the middle of April (Fig. 15c). The rainfall sums over the whole 2 month period measured in Twenthe (30 km northeast of the Hupsel Brook catchment) and Hupsel are similar, leading to similar modelled discharge sums. However, the other two (western) locations experienced up to 60 % more rainfall in Deelen, leading to 100 % more discharge than observed. Between 8 and 11 May 34 mm of rainfall was measured in Deelen, but only 16 mm in Hupsel. This lead to a large overestimation of this discharge peak: 0.15 $mm\,h^{-1}$ in stead of 0.013 $mm\,h^{-1}$ (observed) and 0.033 $mm\,h^{-1}$ (simulated with Hupsel rainfall data).

Precipitation data are the most important forcing data, but not the only ones: observations of potential evapotranspiration, seepage and surface water supply contain errors as well. Potential evapotranspiration estimates obtained at a meteorological station sometimes need preprocessing to become applicable to the whole catchment with its (possible) variety of vegetation. Seepage is difficult to measure and estimates with regional groundwater models are uncertain. Surface water supply is often not measured and modelling decisions of water managers is impossible when changing weir

levels and surface water supply are not automated. In the Cabauw polder, surface water supply was measured, but the uncertainty is large, because the measurement weir was often submerged and because two minor inflow routes were not measured continuously. We estimated the seepage term by closing the water budget for one year and assuming a constant seepage flux year-round. With these assumptions we were able to obtain good results for the Cabauw polder.

In summary, forcing uncertainty is found to be more important than parameter uncertainty and much more important than uncertainty in initial conditions.

7 Conclusions

We tested the newly developed Wageningen Lowland Runoff Simulator (WALRUS, Brauer et al., 2014) for two Dutch catchments: the slightly sloping and freely draining Hupsel Brook catchment and the flat Cabauw polder with controlled water levels. In both catchments, WALRUS performed well, with Nash–Sutcliffe efficiencies of 0.87 (Hupsel) and 0.83 (Cabauw) for the calibration periods and 0.74 (Hupsel) and 0.76 (Cabauw) for the validation period. This limited decrease in performance indicates that the model is not overparameterized.

The model is able to reproduce processes which are important in lowland catchments and explicitly included in WALRUS, such as the groundwater influence on the unsaturated zone, activation of different flowroutes at different stages of catchment wetness, feedbacks between groundwater and surface water, and seepage and surface water supply.

The model was also able to simulate discharge in extremely wet (flash flood in August 2010; NS = 0.64) and dry (summer 1976; NS = 0.84) periods in the Hupsel Brook catchment. Modelled dynamics of groundwater depth, storage deficit and the contribution of quick flow routes are realistic. This indicates that the model is robust and can be used in other climatic conditions than the calibration period and it suggests that the model can also be used to simulate the hydrologic consequences of climate change

(assuming that the parameters are not affected by climate change). In addition, it can possibly be used for early warning of floods and droughts.

The effect of water management operations (varying weir elevations and surface water supply) are also simulated well, owing to the explicit modelling of surface water. This indicates that WALRUS is suitable for catchments that are heavily influenced by human activity, that it can be used to separate the effects of natural processes and human actions on the hydrological variables and that WALRUS is suitable to forecast the effect of different water management practices (scenario analyses).

Comparing modelled catchment effective variables to point-measurement is not trivial. Observed groundwater levels are influenced both by slow and quick flowroutes and should therefore be compared to the (spatially varying) combination of modelled groundwater depth and quickflow reservoir level rather than to the groundwater depth alone.

WALRUS is most sensitive to the wetness index parameter c_W and the groundwater reservoir constant c_G, and to a lesser extent to the quickflow reservoir constant c_Q. The effect of these three parameters could be identified in the discharge times series, which suggests that the model is not overparameterised. The vadose zone relaxation time parameter c_V, however, has a limited effect, cannot be identified in discharge time series alone and may be redundant for most applications. We tested the effect of uncertainty in parameters, initial conditions and forcing and found that the forcing uncertainty was the most important.

In conclusion, the good correspondence between model and observations, identifiability of parameters and computational efficiency are positive characteristics which make WALRUS applicable for research and practice. Recommendations for further research include investigating the possibilities for data assimilation (Liu and Gupta, 2007; Rakovec et al., 2012) and multi-objective calibration (Gupta et al., 1998; Efstratiadis and Koutsoyiannis, 2010), testing the model in catchments with different climates and areas, and regionalisation of model parameters for application in ungauged basins (Merz and Blöschl, 2004).

Acknowledgements. We thank all data providers: KNMI for precipitation and potential evapotranspiration data in the Hupsel Brook catchment (after 1990) and the Cabauw polder, Water Board Rijn and IJssel for the Hupsel Brook discharge data after 2000 and precipitation data from station Wehl, Rijkswaterstaat (1979–1982) and Deltares (August 2010) for soil moisture and groundwater data in the Hupsel Brook catchment, Fred Bosveld (KNMI) for actual evapotranspiration data of the Cabauw polder and Aart Overeem (KNMI) for radar rainfall data. We thank Jantine Bokhorst, Wilco Terink, Matthijs Boersema. Jacques Warmer, Wim Hovius and Marcel Brinkenberg for help with the field work in the Cabauw polder. We thank Olda Rakovec for help with the DELSA parameter sensitivity analysis.

References

Alley, W. M., Healy, R. W., LaBaugh, J. W., and Reilly, T. E.: Flow and storage in groundwater systems, Science, 296, 1985–1990, 2002. 2093

Appels, W. M., Bogaart, P. W., and van der Zee, S. E. A. T. M.: Influence of spatial variations of microtopography and infiltration on surface runoff and field scale hydrological connectivity, Adv. Water Resour., 34, 303–313, 2011. 2093

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R. Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, B. Am. Meteorol. Soc., 82, 2415–2434, 2001. 2099

Beljaars, A. C. M. and Bosveld, F. C.: Cabauw data for the validation of land surface parameterization schemes, J. Climate, 10, 1172–1193, 1997. 2099

Berne, A., ten Heggeler, M., Uijlenhoet, R., Delobbe, L., Dierickx, Ph., and de Wit, M.: A preliminary investigation of radar rainfall estimation in the Ardennes region and a first hydrological application for the Ourthe catchment, Nat. Hazards Earth Syst. Sci., 5, 267–274, doi:10.5194/nhess-5-267-2005, 2005. 2119

Beven, K.: Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, Hydrol. Earth Syst. Sci., 11, 460–467, doi:10.5194/hess-11-460-2007, 2007. 2095

Beven, K. J.: Rainfall–Runoff Modelling: the Primer, 2nd edn., John Wiley & Sons, LTD, Chichester, UK, 2012. 2119
Beven, K. and Freer, J.: A dynamic topmodel, Hydrol. Process., 15, 1993–2011, 2001a. 2096
Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001b. 2103
Bierkens, M. F. P. and Puente, C. E.: Analytically derived runoff models based on rainfall point processes, Water Resour. Res., 26, 2653–2659, 1990. 2096
Bierkens, M. F. P. and van den Hurk, B. J. J. M.: Groundwater convergence as a possible mechanism for multi-year persistence in rainfall, Geophys. Res. Lett., 34, L02402, doi:10.1029/2006GL0283962007, 2007. 2093
Bormann, H. and Elfert, S.: Application of WaSiM-ETH model to Northern German lowland catchments: model performance in relation to catchment characteristics and sensitivity to land use change, Adv. Geosci., 27, 1–10, doi:10.5194/adgeo-27-1-2010, 2010. 2093
Brauer, C. C., Teuling, A. J., Overeem, A., van der Velde, Y., Hazenberg, P., Warmerdam, P. M. M., and Uijlenhoet, R.: Anatomy of extraordinary rainfall and flash flood in a Dutch lowland catchment, Hydrol. Earth Syst. Sci., 15, 1991–2005, doi:10.5194/hess-15-1991-2011, 2011. 2093, 2096, 2097, 2108, 2141
Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., and Uijlenhoet, R.: Investigating storage–discharge relations in a lowland catchment using hydrograph fitting, recession analysis, and soil moisture data, Water Resour. Res., 49, 4257–4264, 2013a. 2108
Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., and Uijlenhoet, R.: The Wageningen Lowland Runoff Simulator (WALRUS): a lumped rainfall-runoff model for catchments with shallow groundwater, Geosci. Model Dev. Discuss., 7, 1357–1411, doi:10.5194/gmdd-7-1357-2014, 2014. 2092, 2093, 2095, 2100, 2102, 2108, 2112, 2117, 2120, 2134
Brooks, R. H. and Corey, A.. T.: Hydraulic Properties of Porous Media, Hydrology Paper 3, Colorado State University, Fort Collins, CO, 27 pp., 1964. 2117, 2147
Chen, T. H., Henderson-Sellers, A., Milly, P. C. D., Pitman, A. J., Beljaars, A. C. M., Polcher, J., Abramopoulos, F., Boone, A., Chang, S., Chen, F., Dai, Y., Desborough, C. E., Dickinson, R. E., Dümenil, L., Ek, M., Garratt, J. R., Gedney, N., Gusev, Y. M., Kim, J., Koster, R., Kowalczyk, E. A., Laval, K., Lean, J., Lettenmaier, D., Liang, X., Mahfouf, J.-F., Mengelkamp, H.-T., Mitchell, K., Nasonova, O. N., Noilhan, J., Robock, A., Rosenzweig, C., Schaake, J., Schlosser, C. A., Schulz, J.-P., Shao, Y., Shmakin, A. B., Verseghy, D. L., Wetzel, P., Wood, E. F., Xue, Y., Yang, Z.-L., and Zengi, Q.: Cabauw experimental results from the Project for Intercomparison of Land-Surface Parameterization Schemes, J. Climate, 10, 1194–1215, 1997. 2099
De Roode, S. R., Bosveld, F. C., and Kroon, P. S.: Dew formation, eddy-correlation latent heat fluxes, and the surface energy imbalance at Cabauw during stable conditions, Bound.-Lay. Meteorol., 135, 369–383, 2010. 2099
Devonec, E. and Barros, A. P.: Exploring the transferability of a land-surface hydrology model, J. Hydrol., 265, 258–282, 2002. 2099
Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, Hydrolog. Sci. J., 55, 58–78, 2010. 2121
Fan, Y., Li, H., and Miguez-Macho, G.: Global patterns of groundwater table depth, Science, 339, 940–943, 2013. 2092
Foken, T.: The energy balance closure problem: an overview, Ecol. Appl., 18, 1351–1367, 2008. 2099
Gilfedder, M., Rassam, D., Stenson, M., Jolly, I., Walker, G., and Littleboy, M.: Incorporating land-use changes and surface–groundwater interactions in a simple catchment water yield model, Environ. Modell. Softw., 38, 62–73, 2012. 2093
Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information, Water Resour. Res., 34, 751–763, 1998. 2121
Gusev, Y. M. and Nasonova, O. N.: The land surface parameterization scheme SWAP: description and partial validation, Global Planet. Change, 19, 63–86, 1998. 2099
Hazenberg, P., Leijnse, H., and Uijlenhoet, R.: Radar rainfall estimation of stratiform winter precipitation in the Belgian Ardennes, Water Resour. Res., 47, W02507, doi:10.1029/2010WR009068, 2011. 2119
Heimovaara, T. J. W. B.: A computer-controlled 36-channel time domain reflectometry system for monitoring soil water contents, Water Resour. Res., 26, 2311–2316, 1990. 2100
Hopmans, J. W. and Stricker, J. N. M.: Stochastic analysis of soil water regime in a watershed, J. Hydrol., 105, 57–84, 1989. 2096
Hopmans, J. and van Immerzeel, C.: Variation in evapotranspiration and capillary rise with changing soil profile characteristics, Agr. Water Manage., 13, 295–305, 1988. 2096

Jacobs, A. F. G., Heusinkveld, B. G., Kruit, R. J. W., and Berkowicz, S. M.: Contribution of dew to the water budget of a grassland area in The Netherlands, Water Resour. Res., 42, W03415, doi:10.1029/2005WR004055, 2006. 2101

Jacobs, A. F. G., Heusinkveld, B. G., and Holtslag, A. A. M.: Eighty years of meteorological observations at Wageningen, the Netherlands: precipitation and evapotranspiration, Int. J. Climatol., 30, 1315–1321, 2010. 2101

Kavetski, D. and Fenicia, F.: Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights, Water Resour. Res., 47, W11511, doi:10.1029/2011WR0107482011, 2011. 2095

Klemeš, V.: Operational testing of hydrological simulation models, Hydrolog. Sci. J., 31, 13–24, 1986. 2095

Koch, S., Bauwe, A., and Lennartz, B.: Application of the SWAT Model for a tile-drained lowland catchment in North-Eastern Germany on subbasin scale, Water Resour. Manag., 27, 791–805, 2013. 2093

Kollet, S. J. and Maxwell, R. M.: Integrated surface–groundwater flow modeling: a free-surface overland flow boundary condition in a parallel groundwater flow model, Adv. Water Resour., 29, 945–958, 2006. 2093

Leijnse, H., Uijlenhoet, R., van de Beek, C. Z., Overeem, A., Otto, T., Unal, C. M. H., Dufournet, Y., Russchenberg, H. W. J., Figueras i Ventura, J., Klein Baltink, H., and Holleman, I.: Precipitation measurement at CESAR, the Netherlands, J. Hydrometeorol., 11, 1322–1329, 2010. 2099

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007. 2121

Makkink, G. F.: Testing the Penman formula by means of lysimeters, Int. J. Water. Eng., 11, 277–288, 1957. 2097, 2099

Maxwell, R. M. and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, Nat. Geosci., 1, 665–669, 2008. 2093

Maxwell, R. M. and Miller, N. L.: Development of a coupled land surface and groundwater model, J. Hydrometeorol., 6, 233–247, 2005. 2093

Melsen, L. A., Teuling, A. J., van Berkum, S. W., Torfs, P. J. J. F., and Uijlenhoet, R.: Catchments as simple dynamical systems: a case study on methods and data requirements for parameter identfication, Water Resour. Res., under review, 2014. 2103

Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, J. Hydrol., 287, 95–123, 2004. 2121

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. 2103

National Institute for Drinking Water Supply: Possible Locations for Deep Groundwater Extraction in West-Utrecht, Tech. rep., 1982 (in Dutch). 2100

NHV: Modelling the Hupsel Brook catchment – contributions from Querner, E., Klutman, W., Droogers, P., Terink, W., Schuphof, A., van Meekeren, B., Verhagen, F., Vermue, H., van der Wal, B., Brauer, C., Kloosterman, P., Teuling, R., Uijlenhoet, R., Pavelková, H., Swierstra, W., Veldhuizen, A., and Willems, G.: Stromingen, in press, 2014 (in Dutch). 2108, 2109

Oreskes, N., Shrader-Frechette, K., and Belitz, K.: Verification, validation, and confirmation of numerical models in the earth sciences, Science, 263, 641–646, 1994. 2095

Overeem, A., Holleman, I., and Buishand, A.: Derivation of a 10-year radar-based climatology of rainfall, J. Appl. Meteorol. Clim., 48, 1448–1463, 2009. 2119

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, Water Resour. Res., 47, W12505, doi:10.1029/2010WR010350, 2011. 2119

Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouweleeuw, B. T., Thielen, J., and de Roo, A. P. J.: Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), Hydrol. Earth Syst. Sci., 9, 381–393, doi:10.5194/hess-9-381-2005, 2005. 2119

Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435–3449, doi:10.5194/hess-16-3435-2012, 2012. 2121

Rakovec, O., Hill, M. C., Clark, M. P., Weerts, A. H., Teuling, A. J., and Uijlenhoet, R.: Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models, Water Resour. Res., 50, 1–18, doi:10.1002/2013WR014063, 2014. 2113, 2145

Refsgaard, J. C. and Knudsen, J.: Operational validation and intercomparison of different types of hydrological models, Water Resour. Res., 32, 2189–2202, 1996. 2095

Rozemeijer, J. C., van der Velde, Y., van Geer, F. C., de Rooij, G. H., Torfs, P. J. J. F., and Broers, H. P.: Improving load estimates for NO_3 and P in surface waters by characterizing

the concentration response to rainfall events, Environ. Sci. Technol., 44, 6305–6312, 2010. 2096
Russchenberg, H., Bosveld, F., Swart, D., ten Brink, H., de Leeuw, G., Uijlenhoet, R., Arbesser-Rastburg, B., van der Marel, H., Ligthart, L., Boers, R., and Apituley, A.: Ground-based atmospheric remote sensing in the Netherlands: European outlook, IEICE Trans. Commun., E88-B, 2252–2258, 2005. 2099
Schuurmans, J. M. and Bierkens, M. F. P.: Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model, Hydrol. Earth Syst. Sci., 11, 677–693, doi:10.5194/hess-11-677-2007, 2007. 2099
Seneviratne, S. I., Lehner, I., Gurtz, J., Teuling, A. J., Lang, H., Moser, U., Grebner, D., Menzel, L., Schroff, K., Vitvar, T., and Zappa, M.: Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event, Water Resour. Res., 48, W06526, doi:10.1029/2011WR011749, 2012. 2110
Stricker, J. N. M. and Brutsaert, W.: Actual evapotranspiration over a summer in the Hupsel Catchment, J. Hydrol., 39, 139–157, 1978. 2096, 2097, 2110
Stricker, J. N. M. and Warmerdam, P. M. M.: Estimation of the water balance in the Hupselse Beek basin over a period of three years and a first effort to simulate the rainfall–runoff process for a complete year, in: Proceedings of the International Symposium on Hydrological Research Basins and their Use in Water Resources Planning, 21–23 September 1982, Bern, Switzerland, 379–388, 1982. 2094, 2096
Tetzlaff, D. and Uhlenbrook, S.: Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data, Hydrol. Earth Syst. Sci., 9, 29–41, doi:10.5194/hess-9-29-2005, 2005. 2119
Teuling, A. J., van Loon, A. F., Seneviratne, S. I., Lehner, I., Aubinet, M., Heinesch, B., Bernhofer, C., Grünwald, T., Prasse, H., and Spank, U.: Evapotranspiration amplifies European summer drought, Geophys. Res. Lett., 40, 2071–2075, doi:10.1002/grl.50495, 2013. 2110
Thom, A. and Oliver, H.: On Penman's equation for estimating regional evaporation, Q. J. Roy. Meteor. Soc., 103, 345–357, 1977. 2097
Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., Prueger, J. H., Starks, P. J., and Wesely, M. L.: Correcting eddy-covariance flux underestimates over a grassland, Agr. Forest Meteorol., 103, 279–300, 2000. 2099

Van den Eertwegh, G. A. P. H.: Water and nutrient budgets at field and regional scale: travel times of drainage water and nutrient loads to surface water, Ph.D. thesis, Wageningen University, 2002. 2096
Van der Velde, Y., Rozemeijer, J. C., de Rooij, G. H., van Geer, F. C., and Broers, H. P.: Field scale measurements for separation of catchment discharge into flow route contributions, Vadose Zone J., 9, 25–35, 2010. 2096
Van der Velde, Y., Rozemeijer, J. C., de Rooij, G. H., van Geer, F. C., Torfs, P. J. J. F., and de Louw, P. G. B.: Improving catchment discharge predictions by inferring flow route contributions from a nested-scale monitoring and model setup, Hydrol. Earth Syst. Sci., 15, 913–930, doi:10.5194/hess-15-913-2011, 2011. 2105
Van der Velde, Y., Torfs, P. J. J. F., van der Zee, S. E. A. T. M., and Uijlenhoet, R.: Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, Water Resour. Res., 48, W06536, doi:2011WR011310, 2012. 2096, 2097
Van Huijgevoort, M. H. J., Hazenberg, P., van Lanen, H. A. J., Teuling, A. J., Clark, D. B., Folwell, S., Gosling, S. N., Hanasaki, N., Heinke, J., Koirala, S., Stacke, T., Voss, F., Sheffield, J., and Uijlenhoet, R.: Global multimodel analysis of drought in runoff for the second half of the twentieth century, J. Hydrometeorol., 14, 1535–1552, 2013. 2110
Van Loon, A. F. and van Lanen, H. A. J.: Making the distinction between water scarcity and drought using an observation-modeling framework, Water Resour. Res., 49, 1483–1502, 2013. 2111
Van Ulden, A. P. and Wieringa, J.: Atmospheric boundary layer research at Cabauw, Bound.-Lay. Meteorol., 78, 39–69, 1996. 2099
Wagener, T.: Evaluation of catchment models, Hydrol. Process., 17, 3375–3378, 2003. 2095
Zambrano-Bigarini, M. and Rojas, R.: A model-independent particle swarm optimisation software for model calibration, Environ. Modell. Softw., 43, 5–25, 2013. 2103
Zampieri, M., Serpetzoglou, E., Anagnostou, E. N., Nikolopoulos, E. I., and Papadopoulos, A.: Improving the representation of river–groundwater interactions in land surface modeling at the regional scale: observational evidence and parameterization applied in the Community Land Model, J. Hydrol., 420, 72–86, 2012. 2093

Table 1. The main catchment characteristics and average annual water budget. f_{XS} denotes surface water supply and f_{XG} seepage (for all abbreviations, see Table A1).

		Hupsel	Cabauw
Size	[km^2]	6.5	0.5
Elevation	[m a.s.l.]	22–35	−1
Slope	[%]	0.8	0
Soil type		0.2–11 m sand on clay	0.7 m clay on peat
Land use: grass	[%]	59	~ 80
maize	[%]	33	~ 15
forest	[%]	3	0
impervious	[%]	5	0
surface water	[%]	1	5
Annual P	[mm]	790	780
ET	[mm]	560	620
Q	[mm]	310	970
f_{XS}	[mm]	0	630
f_{XG}	[mm]	0	100

Table 2. Water balance terms [mm] for the calibration and validation periods. ΔS denotes a change in soil moisture storage – a negative change in soil moisture storage denotes a depletion of the soil reservoir. It is possible for ET_{act} to exceed ET_{pot} in the Hupsel Brook catchment in 1979–1980 because measurements were independent.

	Hupsel				Cabauw			
	cal		val		cal		val	
	obs	mod	obs	mod	obs	mod	obs	mod
$\sum P$	725	–	682	–	723	–	594	–
$\sum ET_{pot}$	587	–	480	–	607	–	635	–
$\sum ET_{act}$	–	531	496	454	574	604	606	629
$\sum Q$	230	249	286	239	668	688	969	1012
$\sum f_{XG}$	0	–	0	–	97	–	96	–
$\sum f_{XS}$	0	–	0	–	359	–	803	–
$\sum f_{GS}$	–	74	–	57	–	22	–	13
$\sum f_{QS}$	–	174	–	189	–	303	–	203
ΔS	–	−54	−15	−11	−62	−110	−92	−143
residual	–	0	−78	0	0	0	10	0

Table 3. Parameter values obtained by optimization with HydroPSO using different objective functions.

fit on:	Q^2	Q	$\sqrt{Q}$	d_G
c_W	400	380	379	107
c_V	1.8	0.8	8.2	0.2
c_G	5.3	5.0	5.0	5.0
c_Q	1	4	12	87

Table A1. Overview of variables, parameters and functions. All fluxes are catchment averages, both external ones (including Q and f_{XS}) and internal fluxes (which are multiplied with the relative surface area of the reservoir in question). Note that d_V, h_Q and h_S result from the mass balances in the three reservoirs, while d_G is only used as pressure head to compute the groundwater drainage flux. The names of the fluxes are derived from the reservoirs (for example f_{XS}: f stands for flow, the X for external and the S for surface water – water flowing from outside the catchment into the surface water network).

	States		
d_V	storage deficit	$\rightarrow \frac{dd_V}{dt} = -\frac{f_{XG}+P_V-ET_V-f_{GS}}{a_G}$	[mm]
d_G	groundwater depth	$\rightarrow \frac{dd_G}{dt} = \frac{d_V-d_{V,eq}}{c_V}$	[mm]
h_Q	level quickflow reservoir	$\rightarrow \frac{dh_Q}{dt} = \frac{P_Q-f_{QS}}{a_G}$	[mm]
h_S	surface water level	$\rightarrow \frac{dh_S}{dt} = \frac{f_{XS}+P_S-ET_S+f_{GS}+f_{QS}-Q}{a_S}$	[mm]
	Dependent variables		
W	wetness index	= func(d_V)	[–]
β	evapotranspiration reduction factor	= func(d_V)	[–]
$d_{V,eq}$	equilibrium storage deficit	= func(d_G)	[mm]
	External fluxes: input		
P	precipitation		[mm h^{-1}]
ET_{pot}	potential evapotranspiration		[mm h^{-1}]
Q_{obs}	discharge (for calibration and Q_0)		[mm h^{-1}]
f_{XG}	seepage (up/down)/extraction		[mm h^{-1}]
f_{XS}	surface water supply/extraction		[mm h^{-1}]
	External fluxes: output		
ET_{act}	actual evapotranspiration	= $ET_V + ET_S$	[mm h^{-1}]
Q	discharge	= func(h_S)	[mm h^{-1}]
	Internal fluxes		
P_S	precipitation into surface water reservoir	= $P \cdot a_S$	[mm h^{-1}]
P_V	precipitation into vadose zone	= $P \cdot (1-W) \cdot a_G$	[mm h^{-1}]
P_Q	precipitation into quickflow reservoir	= $P \cdot W \cdot a_G$	[mm h^{-1}]
ET_V	actual evapotranspiration vadose zone	= $ET_{pot} \cdot \beta \cdot a_G$	[mm h^{-1}]
ET_S	actual evapotranspiration surface water	= $ET_{pot} \cdot a_S$	[mm h^{-1}]

Table A1. Continued.

f_{GS}	groundwater drainage/surface water infiltration	$= \frac{(c_D - d_G - h_S) \cdot \max((c_D - d_G), h_S)}{c_G} \cdot a_G$	[mm h^{-1}]
f_{QS}	quickflow	$= \frac{h_Q}{c_Q} \cdot a_G$	[mm h^{-1}]
	Model parameters		
c_W	wetness index parameter		[mm]
c_V	vadose zone relaxation time		[h]
c_G	groundwater reservoir constant		[mm h]
c_Q	quickflow reservoir constant		[h]
	Supplied parameters		
a_S	surface water area fraction		[–]
a_G	groundwater reservoir area fraction	$= 1 - a_S$	[–]
c_D	channel depth		[mm]
	User-defined functions with defaults		
$W(d_V)$	wetness index	$= \cos\left(\frac{\max(\min(d_V, c_W), 0) \cdot \pi}{c_W}\right) \cdot \frac{1}{2} + \frac{1}{2}$	[–]
$\beta(d_V)$	evapotranspiration reduction factor	$= \frac{1 - \exp[\zeta_1 (d_V - \zeta_2)]}{1 + \exp[\zeta_1 (d_V - \zeta_2)]} \cdot \frac{1}{2} + \frac{1}{2}$	[–]
$d_{V,eq}(d_G)$	equilibrium storage deficit	$= \theta_s \left(d_G - \frac{d_G^{1-1/b}}{(1-\frac{1}{b}) \psi_{ae}^{-1/b}} - \frac{\psi_{ae}}{1-b} \right)$	[mm]
$Q(h_S)$	stage–discharge relation	$= c_S \left(\frac{h_S - h_{S,min}}{c_D - h_{S,min}} \right)^{x_S}$	[mm h^{-1}]
	Parameters for default functions		
ζ_1	curvature ET reduction function		[–]
ζ_2	translation ET reduction function		[mm]
b	pore size distribution parameter		[–]
ψ_{ae}	air entry pressure		[mm]
θ_s	soil moisture content at saturation		[–]
c_S	surface water parameter: bankfull Q		[mm h^{-1}]
x_S	stage–discharge relation exponent		[–]
$h_{S,min}$	surface water level when $Q = 0$		[mm]

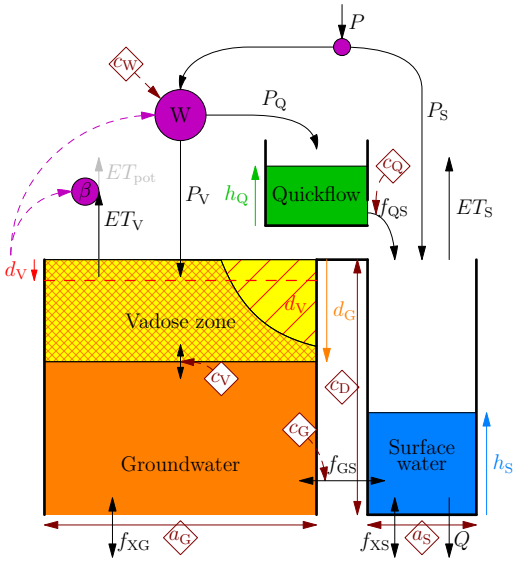

Fig. 1. Overview of the model structure with the five compartments: land surface (purple), vadose zone within the soil reservoir (yellow/red hatched), groundwater zone within the soil reservoir (orange), quickflow reservoir (green) and surface water reservoir (blue). Fluxes are black arrows, model parameters brown diamonds and states in the colour of the reservoir they belong to. For a complete description of all variables, see Table A1 and the accompanying paper (Brauer et al., 2014). The names of the fluxes are derived from the reservoirs (for example f_{XS}: f stands for flow, the X for external and the S for surface water – water flowing from outside the catchment into the surface water network).

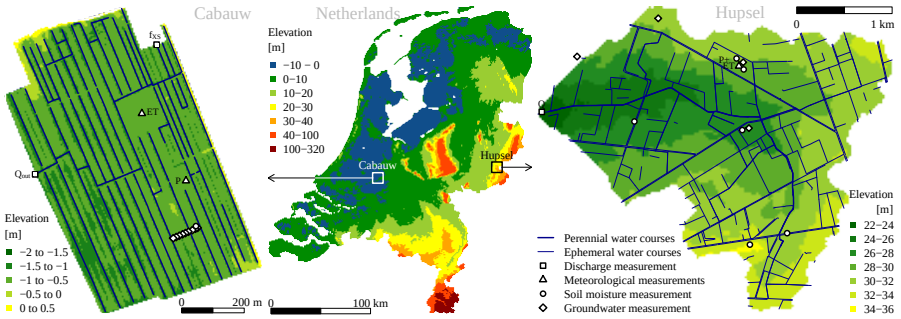

Fig. 2. Elevation maps of the Cabauw polder (left), the Netherlands (middle) and the Hupsel Brook catchment (right) with measurement locations and surface water networks. Soil moisture measurements in the Cabauw polder (circles) consist of 4 arrays of TDR sensors; piezometers (diamonds) are ordered in a transect; f_{XS} denotes surface water supply. In the Hupsel Brook catchment, the circles denote locations of the soil moisture and groundwater observations from the period 1976–1984 and the diamonds denote piezometers used after January 2012.

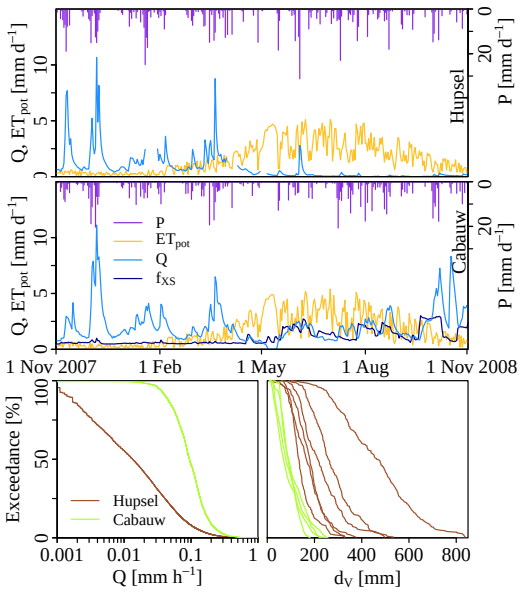

Fig. 3. Example time series of the main water balance terms and regimes of discharge at the catchment outlet Q and storage deficit d_V (i.e. effective thickness of empty soil pores or the volume required to saturate the profile) for the two catchments. Note the effect of surface water supply f_{XS} on the outflow in the Cabauw polder in summer and the discharge regime: discharge is relatively constant throughout the year. The different lines for d_V correspond to different measurement sites, which are well distributed over the Hupsel Brook catchment, but near each other in the Cabauw polder (circles in Fig. 2).

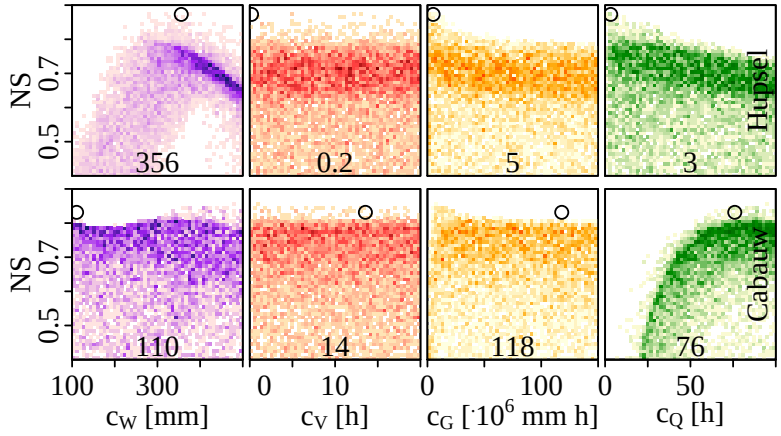

Fig. 4. Relation between Nash–Sutcliffe efficiency and parameter values for The Hupsel Brook catchment (top row) and the Cabauw polder (bottom row). The 10 000 coloured dots are obtained with Monte Carlo analyses. The black circles and numbers indicate the parameter values and resulting Nash–Sutcliffe efficiencies obtained with HydroPSO.

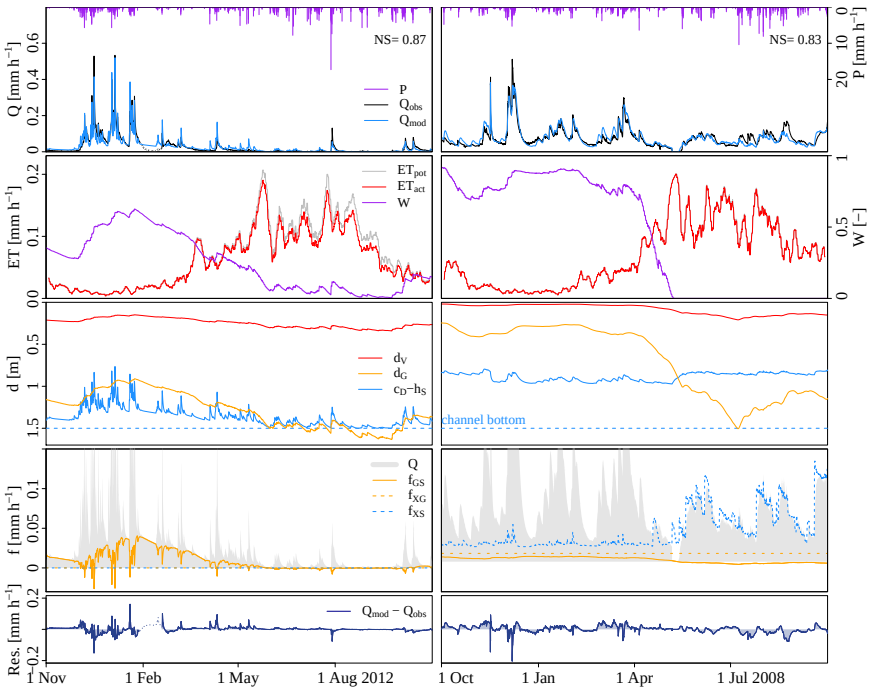

Fig. 5. Model output after calibration. Evapotranspiration data are 5 day moving averages to eliminate daily cycles and focus on long-term differences between ET_{pot} and ET_{act}. The dotted part of Q_{obs} in February 2012 denotes a period with sub-zero temperatures. The surface water level h_S is measured with respect to the channel bottom, while the groundwater depth is measured with respect to the soil surface. The channel depth c_D relates the two to each other. The storage deficit d_V is not a measurable depth, but rather an effective thickness.

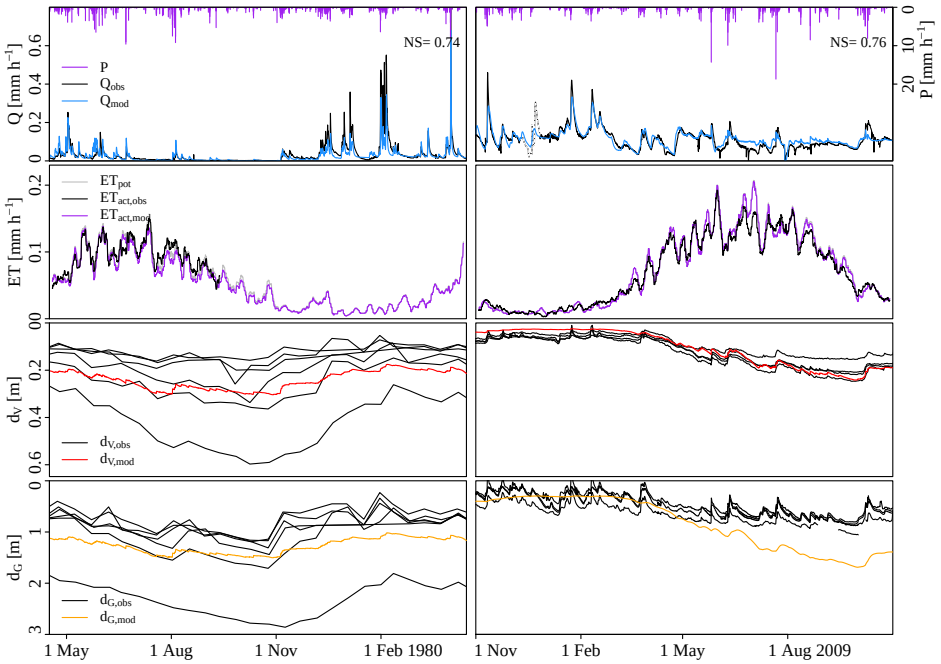

Fig. 6. Validation of model results with discharge, actual evapotranspiration, soil moisture and groundwater data. Note that the temporal resolution of the soil moisture and groundwater data in the Hupsel Brook catchment is 14 days and evapotranspiration data are 5 day moving averages (to eliminate daily cycles and focus on long-term changes). The different lines for observed d_V and d_G represent different locations. A culvert was blocked and opened in the Cabauw polder in December 2008 (dotted part in Q_{obs}).

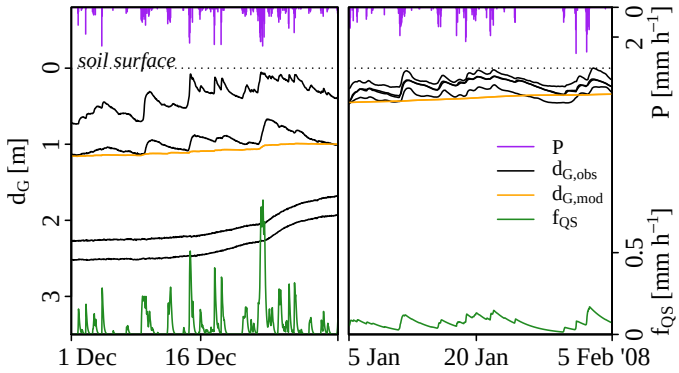

Fig. 7. Comparison of modelled and observed groundwater depths. The different black lines are observations at different locations in the catchments (denoted as diamonds in the map in Fig. 2).

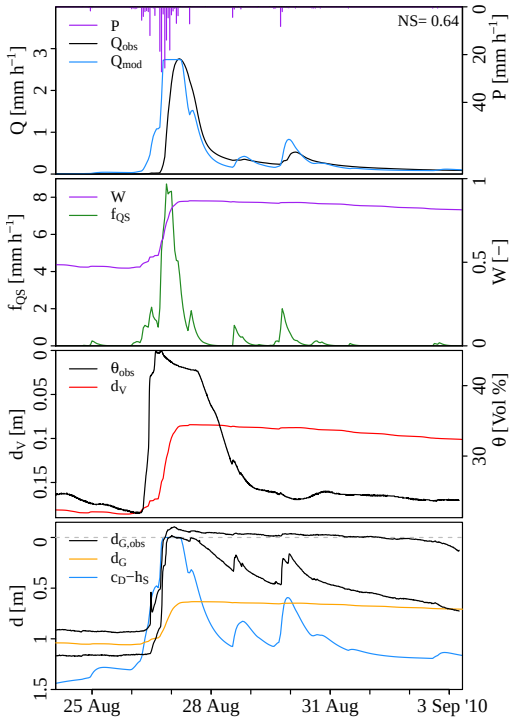

Fig. 8. Simulation of the flash flood in the Hupsel Brook catchment after the extreme rainfall event in August 2010. Groundwater observations were available from two piezometers near the meteorological station: one in a local depression and one in a local elevation (as shown in Brauer et al., 2011). Soil moisture content was measured in the same field (in a local elevation).

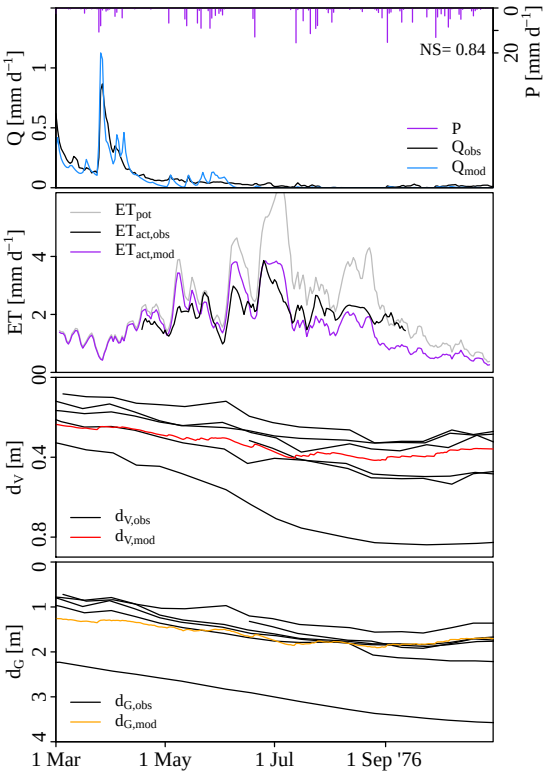

Fig. 9. Simulation of the extremely dry summer of 1976. Note that in contrast to previous model runs, these model output and rainfall and discharge data are daily, groundwater and soil moisture data are biweekly and evapotranspiration are 5 day moving averages.

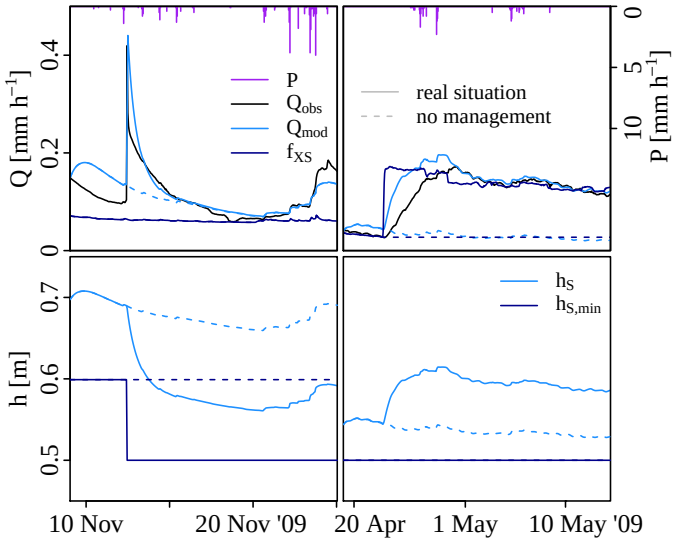

Fig. 10. Reproduction of the Cabauw polder response to water management interventions: weir level decrease (left panels) and increase in surface water supply (right).

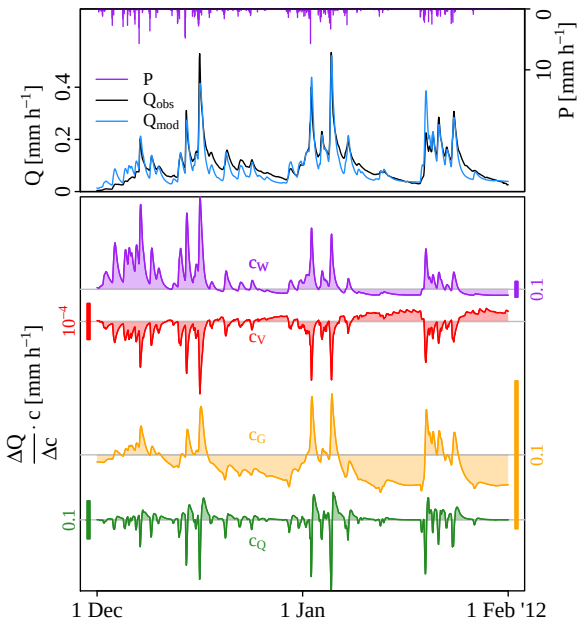

Fig. 11. Identifiability of model parameters in the discharge time series. Top: observed and modelled discharge. Bottom: sensitivity of discharge to a change in each parameter.

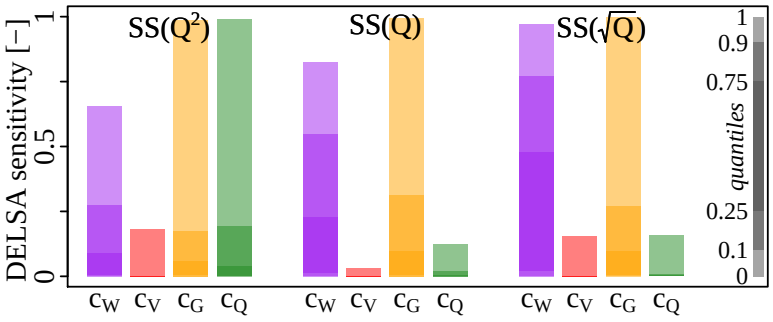

Fig. 12. Parameter sensitivity computed with the Distributed Evaluation of Local Sensitivity Analysis (Rakovec et al., 2014), obtained for three objective functions (SS(Q^2), SS(Q) and SS($\sqrt{Q}$)). The bars show variation between parameter sets as quantiles. Because there are many realisations with low sensitivity, the lower quantiles are zero for most parameters (except c_W for SS(Q) and SS($\sqrt{Q}$)).

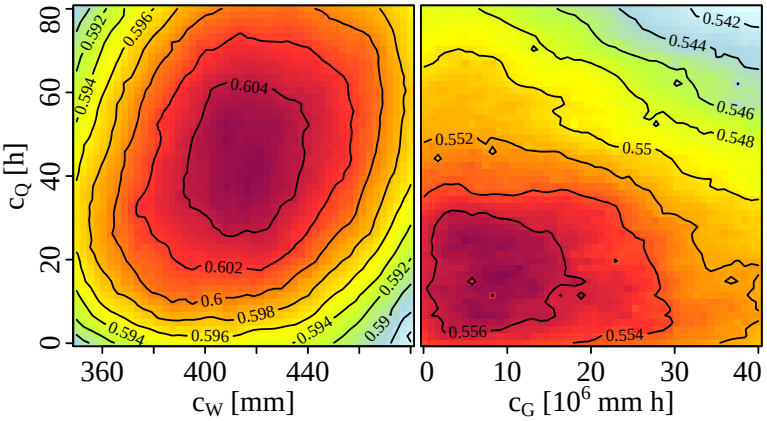

Fig. 13. Examples of response surfaces showing the dependence between parameters. Colours indicate Nash–Sutcliffe efficiencies obtained with Monte Carlo simulations for the Hupsel Brook catchment.

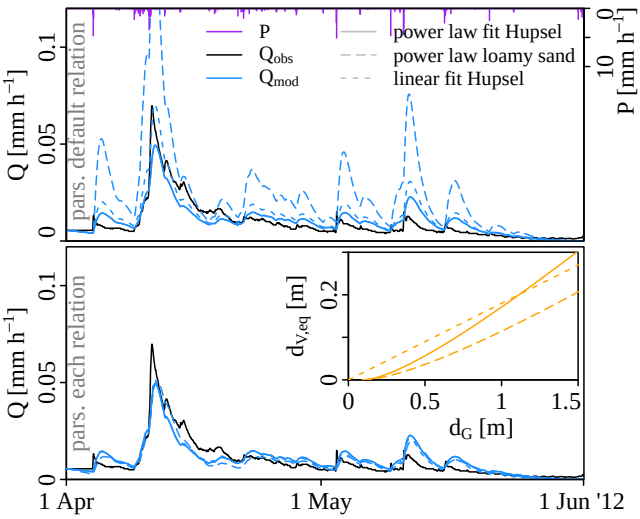

Fig. 14. Effect of the relation between groundwater depth and equilibrium storage deficit. Three options for this relation are plotted in the inset: the relation based on a power-law soil moisture profile (the default), fitted on soil moisture and groundwater observations in the Hupsel Brook catchment (solid; default), the relation based on the theoretical power-law soil moisture profile for loamy sand (long dashed; Brooks and Corey, 1964) and a linear fit between soil moisture and groundwater observations in the Hupsel Brook catchment (dashed).

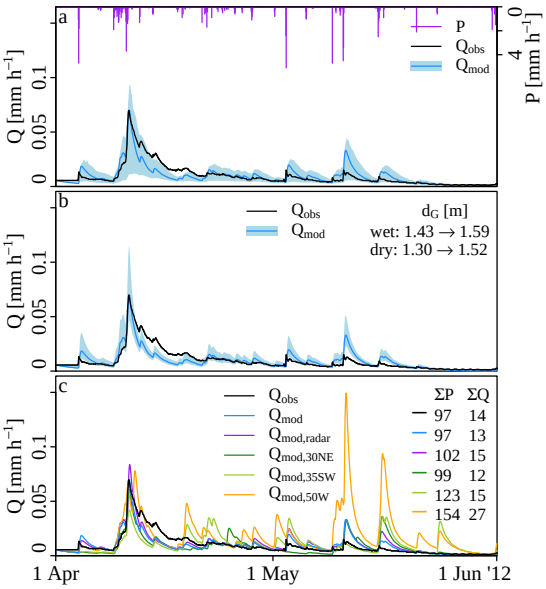

Fig. 15. Propagation of uncertainty in parameters, initial conditions and forcing. **(a)** Range between the 10th and 90th percentile of discharge computed with 100 parameter sets. **(b)** Range between discharges computed with initial groundwater levels based on 0 % and 100 % of discharge originating from drainage. **(c)** Discharge computed with rainfall from radar and rain gauges in the Hupsel Brook catchment, Twenthe (30 km northeast), Wehl (35 km southwest) and Deelen (50 km west).