# Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling

Xin He,[1] Torben O. Sonnenborg,[1] Jens Christian Refsgaard,[1] Flemming Vejen,[2] and Karsten H. Jensen[3]

[1] Weather radar-based quantitative precipitation estimation (QPE) is in principle superior to the areal precipitation estimated by using rain gauge data only, and therefore has become increasingly popular in applications such as hydrological modeling. The present study investigates the potential of using multiannual radar QPE data in coupled surface water—groundwater modeling with emphasis given to the groundwater component. Since the radar QPE is partly dependent on the rain gauge observations, it is necessary to evaluate the impact of rain gauge network density on the quality of the estimated rainfall and subsequently the simulated hydrological responses. A headwater catchment located in western Denmark is chosen as the study site. Two hydrological models are built using the MIKE SHE code, where they have identical model structures expect for the rainfall forcing: one model is based on rain gauge interpolated rainfall, while the other is based on radar QPE which is a combination of both radar and rain gauge information. The two hydrological models are inversely calibrated and then validated against field observations. The model results show that the improvement introduced by using radar QPE data is in fact more obvious to groundwater than to surface water at daily scale. Moreover, substantial negative impact on the simulated hydrological responses is observed due to the cut down in operational rain gauge network between 2006 and 2010. The radar QPE based model demonstrates the added value of the extra information from radar when rain gauge density decreases; however it is not able to sustain the level of model performance preceding the reduction in number of rain gauges.

## 1. Introduction

[2] Precipitation is the main driving force for hydrological processes in most terrestrial water cycles such as in Denmark. From a water resources management perspective, accurate information on the temporal and spatial variation in precipitation plays a critical role in determining the overall water budget at catchment and subcatchment scale [*Watts and Calver*, 1991; *Kuczera and Williams*, 1992; *Boyle et al.*, 2001; *Beven*, 2002; *Younger et al.*, 2009]. The possibility of using precipitation estimated by weather radar for hydrological purposes was initially recognized in the 1970s, when the concept of radar-based quantitative

precipitation estimation (QPE) was introduced to practical hydrological operations such as flood forecasting [*Clark et al.*, 1972; *Brandes*, 1975; *Wilson and Brandes*, 1979]. Since then, an increasing number of research activities have explored the potentials of using radar QPE in hydrological applications [*Carpenter et al.*, 1999; *Tilford et al.*, 2002; *Cole and Moore*, 2008, 2009; *Gourley et al.*, 2010; *He et al.*, 2011b].

[3] The primary advantage of using radar for rainfall estimation is that it provides higher spatial and temporal resolution products and has larger areal coverage compared to traditional rain gauge measured rainfall. This makes radar-based QPE products very suitable for distributed hydrological modeling [*Steiner et al.*, 1995; *Ciach et al.*, 1997; *Gourley and Vieux*, 2005; *Villarini et al.*, 2008; *van de Beek et al.*, 2010]. In principle, radar is able to provide rainfall measurements corresponding to the grid scale of distributed hydrological models whereas rain-gauge-based products are based on geometrical interpolations to each model grid from a limited number of observation points.

[4] Despite the scaling advantages in space and time, stand-alone radar systems may not be able to serve the purpose of quantitative application. This is mainly because radar measures precipitation remotely and indirectly, and the radar signal is often contaminated with various types of

[1] Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark.

[2] Data and Climate Division, Danish Meteorological Institute, Copenhagen, Denmark.

[3] Department of Geography and Geology, University of Copenhagen, Copenhagen, Denmark.

Corresponding author: X. He, Department of Hydrology, Geological Survey of Denmark and Greenland, Oester Voldgade 10, DK-1350 Copenhagen, Denmark. (xhe@geus.dk)

noises and artifacts. As a result, the operational radar QPE is often based on a combination of both radar and rain gauge information [*Barnes*, 1964; *Smith and Krajewski*, 1991; *Anagnostou and Krajewski*, 1998; *Seo and Breidenbach*, 2002; *Haberlandt*, 2007]. However, it is important to know, first, if the rain gauge network density is sufficient to provide adequate bias correction to the radar data; second, if there are the additional information introduced by using radar data in combination with the existing rain gauge network; and third, if the current rain gauge density can be reduced without sacrificing the quality of the rainfall estimate. It may be expected that the relationship between the quality of the radar QPE product would be a nonlinear function of the rain gauge density and that the quality at some point would reach a level where additional rainfall information from extra rain gauges would be redundant [*Krajewski et al.*, 1991; *Abtew et al.*, 1995; *Kim et al.*, 2010]. Solving these issues can be beneficial when evaluating the economic aspects of maintaining the rain gauge network and requires that the rainfall products are tested as driving forces to, e.g., in hydrological models.

[5] Radar precipitation has been used in various types of hydrological studies, where the improvement of simulated hydrological dynamics by using high-resolution data is documented. Among them the most common application of radar QPE data is to simulate or predict flash flood events using rainfall-runoff models [*Carpenter et al.*, 1999; *Ntelekos et al.*, 2006; *Villarini et al.*, 2009; *Atencia et al.*, 2011; *Zoccatelli et al.*, 2011]. However, radar-based precipitation products have never been employed in driving a groundwater model. It is acknowledged in groundwater modeling that the spatial and temporal variation of groundwater recharge plays a significant role in obtaining realistic flow paths and changes in storage [*Bormann et al.*, 1999; *Burns et al.*, 2001; *Liang and Xie*, 2003; *Krause et al.*, 2007; *Milewski et al.*, 2009], and that recharge is highly dependent on precipitation. Therefore, the potential of using radar QPE data in such context may have been overlooked in previous studies. Moreover, the focus of water management at catchment scale has in recent years moved toward using an integrated water resource approach, which requires modeling of the entire water cycle including both the surface and subsurface elements [*Haberlandt et al.*, 2001; *Ludwig et al.*, 2003; *Henriksen et al.*, 2008; *Benito et al.*, 2010]. Considering the high resolution and large spatial coverage, it is likely that using radar QPE data can provide more realistic forcing at the time scales relevant for ground water recharge.

[6] Dynamic, coupled surface water-groundwater models require years of observation data with relatively high quality to obtain reliable water budgets. However, application of long time series of radar QPE data may cause difficulties, since the nonprecipitation echoes as well as other types of errors left on the radar QPE product could accumulate over time and affect the model calibration [*Ryzhkov and Zrnic*, 1995; *Anagnostou et al.*, 1998; *Steiner et al.*, 1999; *Borga et al.*, 2002; *Goudenhoofdt and Delobbe*, 2009; *Villarini and Krajewski*, 2010]. Thus, a thorough model performance evaluation using not only the radar but also rain-gauge-based precipitation product for comparison could help to provide an understanding of the physical processes in the hydrological models. Additionally, the coupled hydrological model is also an indirect tool to assess the accuracy of the rainfall input.

[7] The objectives of the present study are: (1) to prepare multiannual rainfall data sets and to apply both radar and rain-gauge-based rainfall products in hydrological models, (2) to analyze the added value of using higher resolution radar QPE data in hydrological modeling, and (3) to identify the role of rain gauge density on the accuracy of rainfall estimation and hydrological model simulations.

## 2. Study Area and Data

[8] The study area is a catchment located in the western part of Denmark where the Skjern River discharges to the Ringkoebing Fjord (RF), Figure 1. Glacial melt water sand from the last Ice Age constitutes the upper layer of the geology, which enables high infiltration capacity. This feature together with generally shallow groundwater table suggests that the RF catchment may be an ideal place to observe possible interactions between precipitation and groundwater. Two precipitation data sets are prepared in this study, one based on a combination of weather radar data and rain gauge data while the other is based entirely on rain gauge data.

### 2.1. Ringkoebing Fjord Catchment and Observation Data

[9] The RF catchment is a hydrological observatory and the best instrumented area in Denmark for hydrological research activities [*Jensen and Illangasekare*, 2011]. The catchment is approximately 3500 km$^2$, Figure 1. Land use consists mainly of agriculture (85%) while the rest is urban and forest areas.

[10] Western Denmark is dominated by a typical maritime climate. Thus, the RF catchment experiences mild winters, cool summers, and frequent precipitation. On average, precipitation is observed more than a third of the days each year. The mean annual precipitation is estimated to 1057 mm. The catchment is drained by the largest river in Denmark, the Skjern River, which has mean discharge of approximately 35 m$^3$/s.

[11] Observation data are available from stream discharge stations (Figure 1) and groundwater monitoring wells (Figure 2). The observation points are distributed relatively homogenously over the catchment. The groundwater observation data are collected by the local Danish authorities and stored in the Jupiter database maintained by the Geological Survey of Denmark and Greenland (GEUS). Due to an administrative restructuring of the local authorities, many of the groundwater data collected in 2005 and 2006 have not been reported to GEUS. As a result, most of the field data collected during and shortly after 2005–2006 are missing from the database. After 2007, data collection and reporting has been resumed but with a much less dense network as compared to the period before 2005 (seen on Figure 2).

[12] The top soil in the area is highly permeable and the stream discharge is therefore dominated by inflow from groundwater (base flow). In large areas, the groundwater level is close to ground surface and subsurface drainage pipes are therefore installed to direct the excessive water to the streams. The shallow geology is
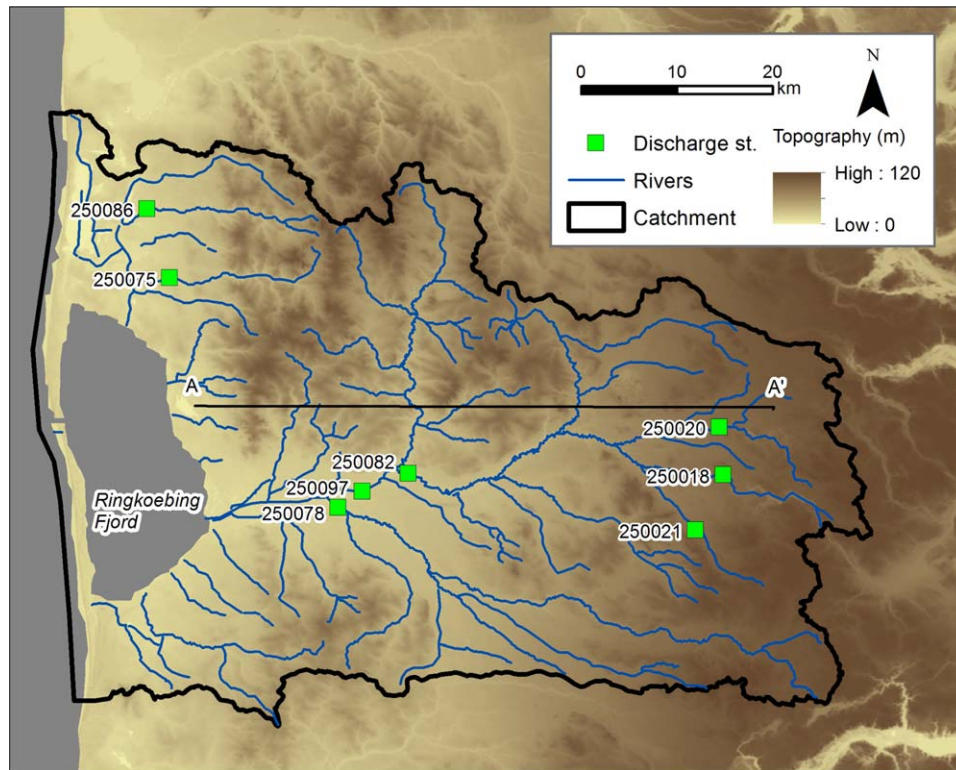
**Figure 1.** Hydrological model domain for the Ringkoebing Fjord (RF) model. Skjern River is the main river branch in the catchment. However, smaller stream discharging to the Ringkoebing Fjord are also included in the model. A cross section (A–A′) of the geological structure is shown in Figure 3.

dominated by Quaternary sand and gravel which forms large interconnected aquifers. Below, large Miocene sand and clay formations are found. Data from groundwater abstraction wells are the main data source for building the geological model. Six geological units were defined,

three Quaternary units comprising fractured clay located close to the soil surface, sand and clay, together with three pre-Quaternary units defined as clay, Miocene mica sand and Miocene quartz sand. A geological cross section is shown in Figure 3.
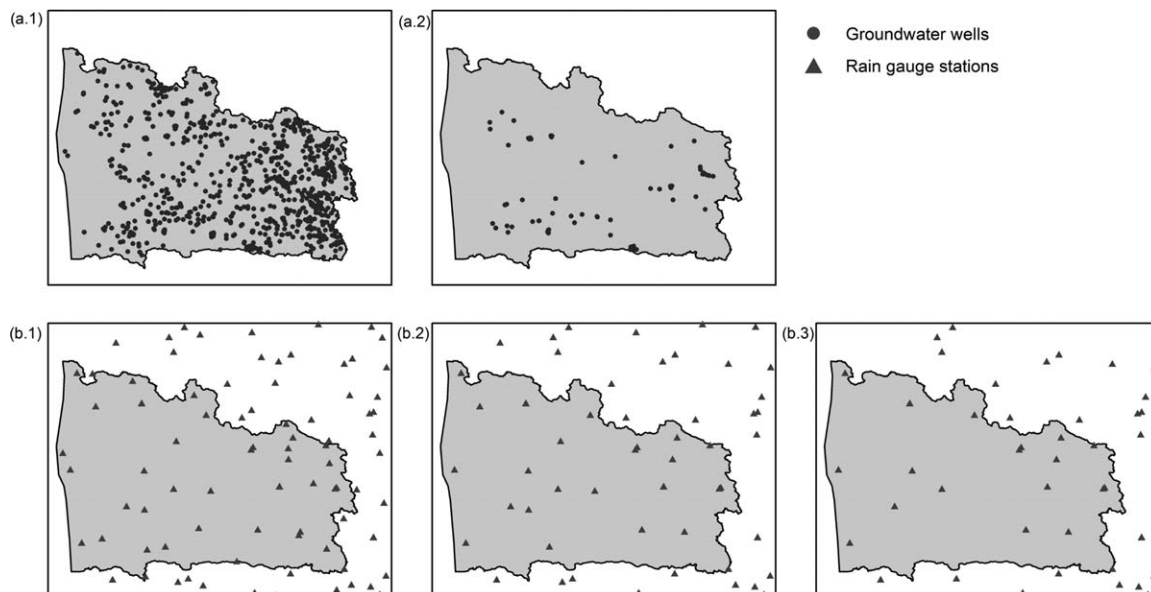


**Figure 2.** Change of observation network density in Skjern catchment: (a) Groundwater head observation network with (a.1) the status in 2005 and (a.2) the status in 2007; (b) the rain gage network for estimating precipitation with (b.1) the status in 2006, (b.2) the status in 2009, and (b.3) the status in 2010.
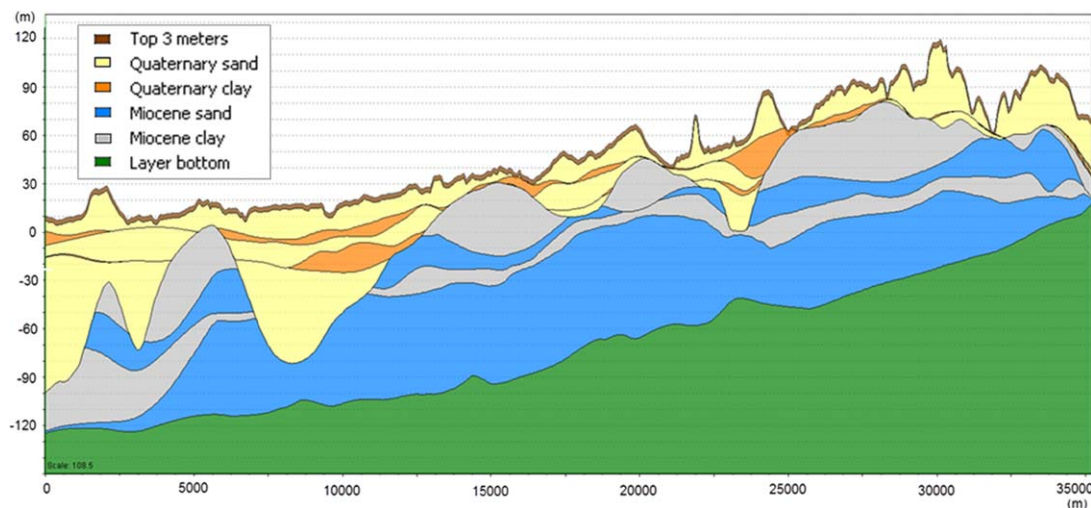
**Figure 3.** Cross-section of the geological structure from west to east in the RF catchment. Location of the transect is shown in Figure 2.

## 2.2. Radar-Based Precipitation

[13] The radar data used in the study are retrieved from two C-band radars located at Roemoe and Sindal in western Denmark as shown on Figure 4. The radars are operated by The Danish Meteorological Institute (DMI) and have a detection range of 240 km. Pseudo-CAPPI images of radar reflectivity at 2 km height are generated every 10 min. A maximum-pixel-value approach is used to generate composite sums in order to avoid the borderline effect in the area where the radar beams intercept.

[14] The radar QPE algorithm used in this study is developed by DMI where the basic principles of the algorithm are based on the method developed by [*Michelson and Koistinen*, 2000]. The program runs on hourly or daily time step and the end product, short name ARNE, has a resolution of 2 km Cartesian grid. Detailed description of the ARNE algorithm can be found in [*He et al.*, 2011a], and a brief summary is given in the following.

[15] The radar reflectivity factor (Z) is converted to rain rate (R) using the power law Marshall-Palmer relationship [*Marshall et al.*, 1947]. After the hourly accumulated rainfall is obtained, the raw radar images are subjected to bias adjustment based on ground observations from rain gauges. The radar pixels with collocated rain gauges are collected as gauge-radar data pairs. The bias is expressed as the logarithmic transformation of the gauge-to-radar ratio. After interpolation of the bias values to the entire grid domain, the final adjusted daily radar precipitation is obtained by multiplying an adjustment factor field to the original radar composite sum. As such, ARNE is a range-dependent bias adjustment algorithm that not only adjusts the mean bias but also conditions individual pixel values based on the location of the rain gauges. However, the method for conditioning radar rainfall field to rain gauge records is not an exact procedure. The rain gauge data are used for establishing the bias-to-distance relationship, thus the rain gauge records at their corresponding locations are not preserved to the exact values.

[16] When the radar QPE images are accumulated, static ground clutter are magnified since no clutter removal program has been implemented by the radar operator. These accumulated clutters can cause large errors in the hydrological model since they are pure noises. Thus, in the present study the static clutters are removed by using a simple clutter filter. Histograms are generated using all pixels in the annual accumulated radar images. Under the assumption that an accumulated clutter signal is stronger than a precipitation signal, the locations of the pixels at the top 2% of the histogram are identified and the values are recalculated using the average values from adjacent pixels. However, the procedure is not able to deal with neither dynamic



**Figure 4.** Location of the study area (shaded) in the western part of Denmark. The Danish national radar network used in the radar-based QPE algorithm includes the radars at Roemoe, Sindal, and Stevns. The detection range is shown by both 120 km and 240 km. The study area can be reached by all three radars but is mostly influenced by the Roemoe radar.

clutters that change in space and time nor the static clutters with signal strength similar to the rainfall signal.

### 2.3. Rain Gauge Data

[17] Ground-based precipitation data are obtained from the rain gauges located across the RF catchment. In the present study, the 10 km rainfall grid product provided by the Danish Meteorological Institute (DMI), DMI10 for short in the following will be used as a comparison to the radar-based rainfall product. Since many of the rain gauge stations in the study area are manual gauges with daily observation frequency, both radar and rain-gauge-based precipitation products used in the hydrological model represent accumulated daily precipitation. Inverse distance weighting interpolation is used to project point observation data onto the 10 km rectangular grids. Instead of using a fixed search radius, where all stations within a defined distance are included, the method seeks in the four sectors around the center point, and uses only the nearest station in each sector. The stations used in this method are always the best available in terms of geographical spreading [*Scharling et al.*, 2006].

[18] In principle, the radar image adjustment described in section 2.2 takes account of all the rain gauges in the country, but since the multiplicative factor is reversely proportional to the distance, it is estimated that rain gauges located further than 50 km away will have little or no contribution to the adjustment of a particular pixel. It is difficult, however, to identify the exact number of rain gauges that affect the resulting radar images of the RF catchment area. Therefore, a 100 km × 80 km rectangle is selected to demonstrate the spatial layout of rain gauges both within and in the vicinity of the catchment. As seen on Figure 2, the number of rain gauges has been reduced over the country and over the catchment due to system upgrade and budget cuts. In the year of 2006 where the rain gauge number is at its peak, there are 87 gauges inside the rectangle. However, in 2009 and 2010 the number of operational gauges in the catchment is 49 and 40, respectively. Hence, the gauge density has been reduced by more than a factor of two.

[19] The DMI10 product has long been considered as the "standard" rainfall data set in Denmark, and most of the previous hydrological models that use gridded rainfall data as model input have employed such data [*van Roosmalen et al.*, 2009; *Fu et al.*, 2011; *Seifert et al.*, 2012; *Stisen et al.*, 2012].

[20] Both the radar and rain-gauge-based rainfall products are highly dependent on the accuracy of the rain gauge data. Thus, bias correction of the rain gauge observations is crucial to the success of the hydrological modeling. The rain gauge catch correction method proposed by *Stisen et al.* [2011, 2012] is used in this study where air temperature, wind speed, and rainfall intensity on a daily basis at grid scale are accounted for.

## 3. Methods

[21] Our intention is to construct two hydrological configurations where different rainfall estimation schemes can be tested. To make sure these two models are both valid for their respective conditions, the two models also need to be calibrated against observation data. As illustrated in Figure 5, the overall simulation period is 1996–2010 with the first 10 years used as a warm-up period. The DMI10 product is used during the warm-up period to ensure identical hydrological conditions at the beginning of 2006 for the different rainfall scenarios introduced in the later period. Radar QPE data are prepared for 2006–2010. Due to the reduction in rain gauges, model calibrations are carried out for the period 2007–2009 and the calibrated models are validated in 2006 and 2010. In this way, both the effect of rain gauge density and the impact of alternative rainfall products on the hydrological model performance are evaluated. When designing the experiment shown in Figure 5, a number of considerations have been given regarding the data discontinuity. Rain gauge data are available for the entire simulation period (1996–2010), whereas the radar data are only available from 2006 to 2010. Therefore, the discontinuity problem does not exist in the rain-gauge-based model. In the radar-based model, a switch of rainfall product takes place at 2005/2006. However, both years have the "full" rain gauge network, which makes the bias level staying basically the same. Thus, one would expect that radar QPE gives similar rainfall mean as the rain gauge rainfall if radar data before 2006 had existed. Moreover, the model calibration starts at 2007, which means there is one extra year for the groundwater stage to reach equilibrium again even if a difference in the mean value between the rainfall products was present.
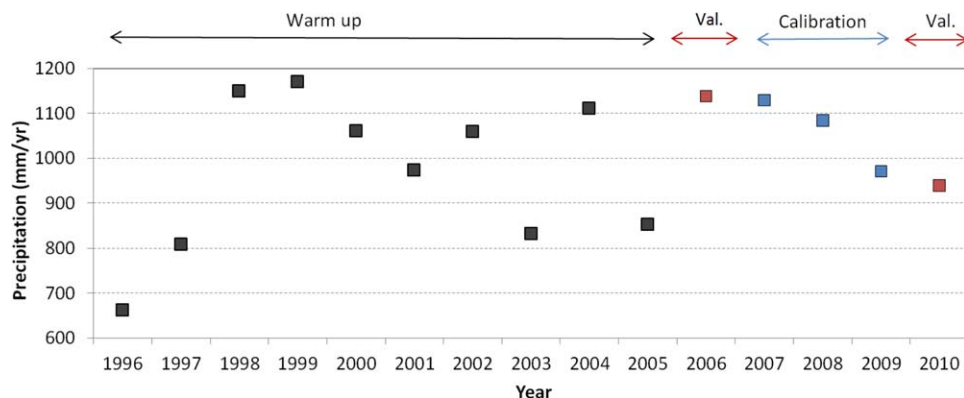


**Figure 5.** Annual mean precipitation estimated by rain gauges in the RF catchment. The hydrological model experimental design is shown at the top of the data series.

## 3.1. Hydrological Model and Model Setup

[22] The hydrological model for the RF catchment is developed using the MIKE SHE code which is a deterministic, distributed and physically based modeling system [*Abbott et al.*, 1986] providing a fully coupled groundwater-surface water description. The basic setup of the RF model is in line with the Danish National Water Resources Model, the DK-model, which operates at daily temporal scale on a 500 m grid domain [*Henriksen et al.*, 2003; *Hojberg et al.*, 2012; *Stisen et al.*, 2012].

[23] The current model configuration includes a 2-D diffusive wave approximation of the Saint Venant equations for overland flow; a kinematic routing method for flow in the river system; a two-layer water balance method for the unsaturated zone describing the distribution of infiltration between evapotranspiration and groundwater recharge; a linear reservoir model for flow in subsurface drains; a 3-D Boussinesq equation for flow in the saturated zone; a Darcy flow method that includes both river bed and aquifer conductance for the river-aquifer interaction; and a degree-day approach for snow accumulation and melting. More information about the MIKE SHE modeling system and the descriptions of the involved hydrological processes can be found in *Abbott et al.* [1986], *Refsgaard and Storm* [1995], and *Graham and Butts* [2006].

[24] As the backbone of the groundwater model component in MIKE SHE, a geological model is established for the catchment. The geological model consists of 17 geological units where the hydraulic conductivity, specific yield, and specific storage values are specified for each unit. To accelerate the numerical computation, 11 computational layers are specified which delineate the main aquifers and aquitards. A detailed description of the geological setup of the RF catchment model can be found in *Nyegaard et al.* [2010].

## 3.2 Model Optimization

[25] The model calibration scheme is in accordance with the guidelines given by *Madsen* [2003], where the whole optimization process can be divided into three key steps: The first step is model parameterization where the sensitivities of the model parameters are analyzed and choices of parameters for calibration are made. In the second step, calibration criteria are defined and the objective functions are formulated. The algorithm used for calibration is specified in the third step.

[26] Due to the complexity of the MIKE SHE model structure, multiobjective calibration is adopted which ena-bles various hydrological processes to be optimized simultaneously [*Gupta et al.*, 1999]. Using the multiobjective calibration also prevents overconditioning to a particular water compartment due to limited field data. This feature is very beneficial in the current study since the entire hydrological cycle is considered altogether.

[27] Model calibration is carried out against observation data for hydraulic head and stream discharge, where PEST [*Doherty*, 2003] is used as the optimization tool. PEST has been intensively used in the field of hydrological research [*Moore and Doherty*, 2005; *Skahill and Doherty*, 2006; *Christensen and Doherty*, 2008]. The model optimization method used by PEST is the Levenberg-Marquardt Algorithm (LMA) with Jacobian matrix updating, which is a gradient-based local parameter search algorithm.

[28] The RF model used for the present study has 96 parameters that can be individually specified. Based on a sensitivity analysis, a subset of nine parameters was selected for optimization. Another 14 parameters were tied to the nine free parameters. The rest of the parameter values were fixed at values specified based on previous model applications in the catchment. To minimize the number of free parameters, it is assumed that the ratios between the summer and winter root depths are fixed, and the ratios between the root depths for different soil types are also fixed.

[29] All free parameters are assumed to be log-normally distributed. No constrains were specified to limit the parameter search, implying that the optimized parameter values might end up with unrealistic numbers. This specification was purposely designed since unrealistic parameter values indicate that errors may have occurred in the model input or the model formulation [*Stisen et al.*, 2011]. The selected free parameters as well as the initial values are summarized in Table 1. The initial parameter values are obtained from calibration of a larger scale hydrological model that covers the entire mid-Jutland, including the RF catchment [*Stisen et al.*, 2012].

## 3.3. Model Performance Evaluation

[30] Six objective functions are defined for both model calibration and independent model validation. The objective functions are listed in Table 2 along with their mathematical expressions and the number of observation data in each group. As seen from Table 2, groundwater is more intensively sampled than surface water. Therefore, weighting factors in the multiobjective calibration are assigned to

**Table 1.** List of Parameters As Well As Their Initial Values Used for Calibrating the Skjern Model

| Symbol | Parameter Description | Initial Value | Unit |
|---|---|---|---|
| Kx_1 | Hydraulic conductivity for Quaternary sand | $3.94 \times 10^{-4}$ | m/s |
| Kx_2 | Hydraulic conductivity for Quaternary clay | $3.42 \times 10^{-7}$ | m/s |
| Kx_3 | Hydraulic conductivity for Miocene quartz sand | $8.40 \times 10^{-4}$ | m/s |
| Kx_4 | Hydraulic conductivity for Miocene mica sand | $9.27 \times 10^{-5}$ | m/s |
| Kx_5 | Hydraulic conductivity for pre-Quaternary clay | $6.73 \times 10^{-8}$ | m/s |
| Kx_6 | Hydraulic conductivity for Top 3 m till/moraine | $1.48 \times 10^{-4}$ | m/s |
| Drain | Time constant for drain flow | $2.86 \times 10^{-8}$ | 1/s |
| Leak | River-aquifer leakage coefficient | $1.38 \times 10^{-6}$ | m/s |
| RD | Summer root depth | 600 | mm |

**Table 2.** List of Objective Functions Used for Calibrating the Skjern Catchment Model With PEST

| Objective Function | Equation | Unit | Number of Observations |
|---|---|---|---|
| Nash-Sutcliffe efficiency of daily discharge | $E_i = \dfrac{\sum_{j=1}^{n_i} \left(\Psi_{obs,j} - \overline{\Psi}_{obs,i}\right) - \sum_{j=1}^{n_i} \left(\Psi_{obs,j} - \overline{\Psi}_{sim,i}\right)}{\sum_{j=1}^{n_i} \left(\Psi_{obs,j} - \overline{\Psi}_{obs,i}\right)}$ | – | 10 |
| Mean relative error in the total water balance | $F_{bal,i} = \dfrac{\overline{Q}_{obs,i} - \overline{Q}_{sim,i}}{\overline{Q}_{obs,i}}$ | % | 10 |
| Mean relative error in the summer water balance | | % | 10 |
| RMSE of the individual hydraulic head time series | $RMS_i = \sqrt{\dfrac{1}{n_i} \cdot \sum_{j=1}^{n_i} \left(\Psi_{obs,j} - \Psi_{sim,j}\right)^2}$ | m | 236 |
| RMSE of the mean hydraulic heads in 2000–2005 | | m | 832 |
| Mean error of hydraulic head in each model layer | $ME_i = \dfrac{1}{n_i} \cdot \sum_{j=1}^{n_i} \left(\Psi_{obs,j} - \Psi_{sim,j}\right)$ | m | 9 |

each observation data to make sure the aggregated relative weights among the objective functions are equal.

[31] Summer water balance is introduced as an objective function to adjust the simulated low flows during the dry season more rigidly, whereas the simulated water balance during the wet season is rather reliable based on our past experiences. Because of the groundwater head data availability issue explained before, mean hydraulic head data from the period 2000–2005 where more abundant observation data are available, are also used for calibration. The change in the dynamics of simulated groundwater head is captured by using time series of hydraulic head data from the fewer observation wells in 2007–2009.

## 4. Results

[32] Since the resolution of the precipitation products (2 and 10 km) are coarser than the grid size of the hydrological model (500 m), the grid size of the precipitation estimates are split into 500 m grids to match with the grid size of the RF model without any interpolation. When the precipitation input is switched between radar and rain gauge products, the remaining parts of the RF model setup are unchanged. Therefore, as the hydrological models are calibrated individually against field observations, the differences in the calibrated parameter values only reflect the influence of the rainfall forcing. Based on the calibrated models, validations are carried out with respect to surface water, groundwater, and the overall water balance.

### 4.1. Precipitation Estimated by Radar and Rain Gauge

[33] Although the radar QPE algorithm uses rain gauge data for bias adjustment, it is still necessary to scrutinize the residual bias. One effective and commonly used method for this purpose is to calculate the mean field bias (MFB) between two images [*Smith and Krajewski*, 1991; *Anagnostou et al.*, 1998; *Seo et al.*, 1999; *Borga et al.*, 2002; *Goudenhoofdt and Delobbe*, 2009]. Mean field bias is defined as: $MFB = \sum_{i=1}^{n} G_i \bigg/ \sum_{i=1}^{n} R_i$

[34] where $G_i$ is rain gauge observations over a certain area and $R_i$ is the radar estimated values at the pixels that contain $G_i$. In the present study, daily MFB is calculated for collocated rain gauge-radar data pairs with radar-based QPE as the denominator. The data set is screened for precipitation values larger than 0.3 mm/day in order to reduce the background noises. The data set is also log-transformed for easy visualization.

[35] As seen in Figure 6, the distribution of the log-transformed residual bias from the radar QPE follows the shape of the normal distribution. This confirms that the radar rainfall bias adjustment needs to be operated in a log-based domain which is one of the assumptions in the ARNE algorithm. The bell cloud peaks sharply at zero indicating that the majority of the mean bias has been removed by the ARNE algorithm. On the other hand, it is neither practical nor necessary to remove all the mean bias considering the different nature of these two rainfall products. Toward the lower precipitation intensities, it is observed that the differences between the two estimates can be as large as a factor of 10, which are typically caused by radar image
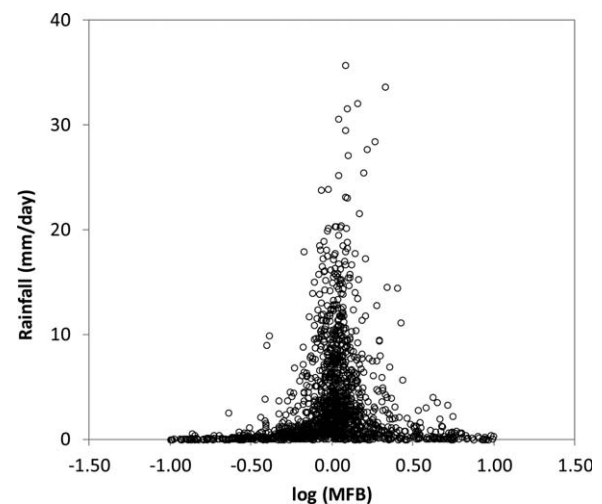


**Figure 6.** Distribution of daily logarithmic mean field bias (MFB) in the RF catchment with different rainfall intensity. Data collected during 2006–2010 with daily mean precipitation larger than 0.03 mm/day.

artifacts. The radar image artifacts can possibly originate from both meteorological and nonmeteorological echoes; however, in the present case, it is mainly due to clutters in the study area. According to DMI, the bright band contamination is very seldom in western Denmark and the VPR correction is not applied. Overall, for the annual water balance the simulated catchment flows are essentially controlled by large rainfall events. As heavy rainfall occurs very rarely in the RF catchment, the overall result of the bias removal is acceptable.
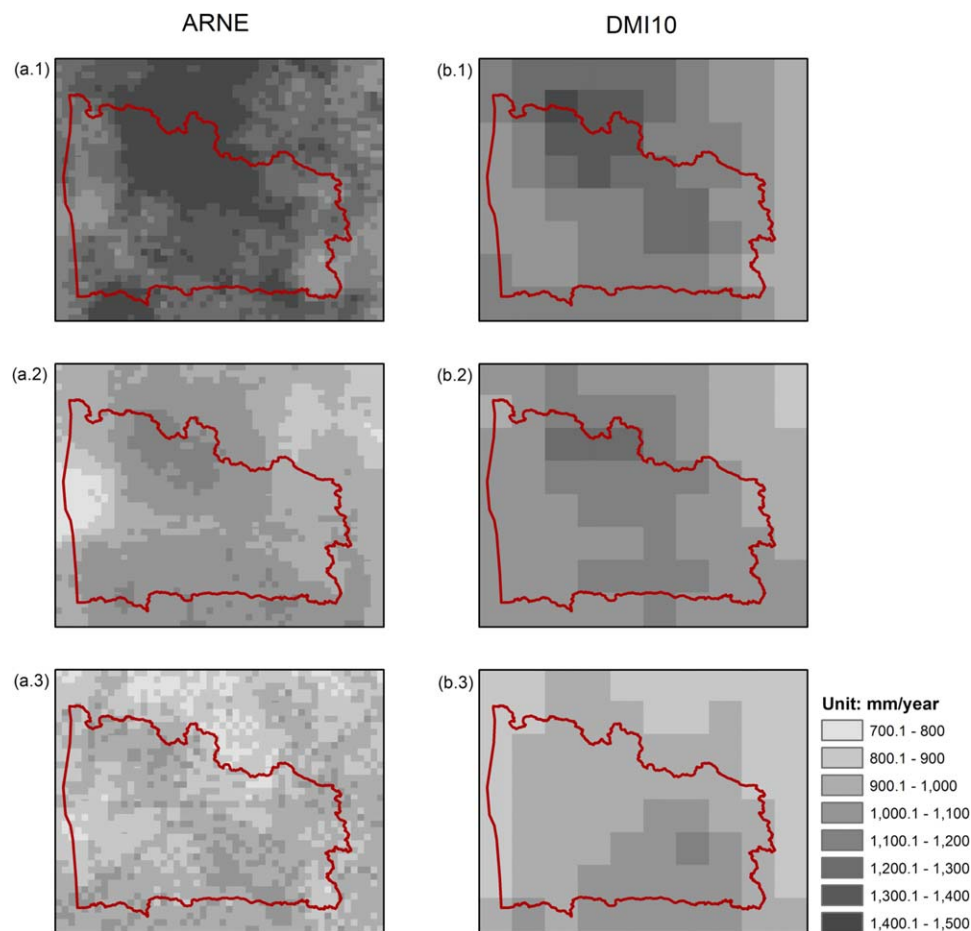
[36] The areal precipitation in the RF catchment for both radar and rain-gauge-based precipitation products are illustrated in Figure 7. Annual averaged rainfall maps are calculated for three periods: 2006, 2007–2009 and 2010, which are consistent with the model calibration and validation periods. The radar-based precipitation product has a five-time higher spatial resolution and is obviously more pixelated than the rain-gauge-based product.

[37] Apart from the mean bias, another common problem of using radar QPE in a hydrological context is that the bias is a function of the distance from the radar site. There are two main factors that could cause radar rainfall estimates to show biases with range: First, the height of the radar beams increase with range, and thus intercept with the vertical profile of reflectivity at different heights, and second, the sampling size of rainfall observation increases with range,

as the radar beams become wider [*Zawadzki*, 1984; *Kitchen et al.*, 1994; *Ryzhkov and Zrnic*, 1995; *Borga and Tonelli*, 2000; *Krajewski and Vignal*, 2001; *van de Beek et al.*, 2010; *Villarini and Krajewski*, 2010]. For this reason, a second order polynomial is established to correct the bias in the ARNE algorithm based on a distance relationship. It is seen in Figure 7 that after the adjustment, the spatial patterns of the two rainfall products are more similar than if only the mean bias is removed from the accumulated radar image [*He et al.*, 2011a], which indicates that the distance induced bias is basically removed from the radar rainfall product and the annual water budget calculated by these two products will be mainly dependent on the mean values inside the study area.

### 4.2. Optimized Hydrological Model Parameters

[38] Verification of radar-based rainfall estimation against rain gauge data alone may be insufficient. This is because the large differences observed between the two data sources are likely to be representativeness errors due to the different supporting scales of the two methods. In the present study, a distributed hydrological model with spatially distributed model outputs serves as a tool for additional verification of the rainfall signal. Before forward model simulations are carried out, model parameter optimizations are performed using DMI10 and ARNE as



**Figure 7.** Radar and rain-gauge-based precipitation estimation products for the RF catchment: (1) average in 2006, (2) average in 2007–2009, and (3) average in 2010.

precipitation input while the rest of the model setup is unchanged. Nine parameters are calibrated using the initial values listed in Table 1 and the objective functions in Table 2.

[39] The nine calibration parameters selected can be grouped into three categories: (1) hydraulic conductivities controlling groundwater flow; (2) drainage and leakage coefficients controlling flow contributions to stream discharge; and (3) root depth controlling evapotranspiration and percolation out of the root zone. The resulting estimates of the optimized parameters are shown in Figure 8. The hydraulic conductivities for the subsurface sediments are very similar for the two model optimizations. The hydraulic conductivity of the top soil ($Kx\_6$) controls the dynamics of the groundwater flow in case the groundwater table is close to the ground surface. In the RF catchment more than half of the area has a groundwater table at less than 3 m depth, which makes the conductivity of this geological layer sensitive to changes in precipitation at daily time scale. Therefore, the difference between the estimates for $Kx\_6$ is relatively larger than for other hydraulic conductivities. Since ARNE gives higher values of estimated precipitation amount than DMI10 during 2007–2009, it is not surprising that the largest difference is seen for the root depth as actual evapotranspiration is closely related to root depth. The calibrated root depth based on ARNE input is about 40% higher than when using DMI10 in order to obtain acceptable overall water balance.

[40] In general, the model parameters that cover a wide range of hydrological processes do not show any significant differences between the two methods.

## 4.3 Model Calibration and Validation Results

[41] Independent validation is essential in hydrological model development and application, since the model parameters may be overconditioned during the calibration period. In that case, the model cannot generate reliable simulations for periods or situations that are different from those in the calibration period. Moreover, results from the hydrological model can be seen as an indicator of the quality of the rainfall products. For these reasons, model performances are evaluated using the objective functions

shown in Table 2 for the two validation periods, 2006 and 2010, individually. Based on the field observations available, the model validations are performed for stream discharge, groundwater head, and the total water balance. Scores of the objective functions from the calibration and validation periods are presented side by side for comparison, Table 3–5.

### 4.3.1. Stream Discharge

[42] Nash-Sutcliffe efficiency (NSE) and two water balance scores are shown in Table 3. The discharge stations are sorted by catchment size where the locations can be found in Figure 1. During the calibration period (2007–2009), the radar-based model is superior to the rain-gauge-based model in terms of NSE. However, in the same period the rain-gauge-based model is more favorable in terms of water balance scores showing smaller bias for most stations both with respect to annual accumulation and accumulation over the summer months (June, July, and August). This indicates that ARNE generates a larger mean bias but a smaller residual variance over the calibration periods. For stream discharge we consider NSE to be more powerful for model evaluation since it accounts for the timing of the peaks and also the shapes of the recession curves. Thus, the radar-based model is considered more accurate during calibration in this case. In Figure 9, simulated hydrographs at an upstream station (250021) and a downstream station (250097) are shown. Based on visual inspection, the differences between the model simulations are larger at the upstream site, however, the difference between the Nash-Sutcliffe scores are similar for the two stations in the calibration period 2007–2009 with values of 0.73 and 0.67 at the upstream station, and 0.77 and 0.72 at the downstream station. This trend is also obtained at other upstream stations not included in the figure. However, both models underestimate discharge in the second half of 2008, which explains the larger error in the water balance scores. It is also noted that during winter between 2006 and 2007, both models produce discharge that are much higher than observed. This problem is presumably caused by a model deficiency in the two-layer description of the unsaturated zone. Before the large peaks 10 successive days with significant rainfall were observed. Since the unsaturated zone
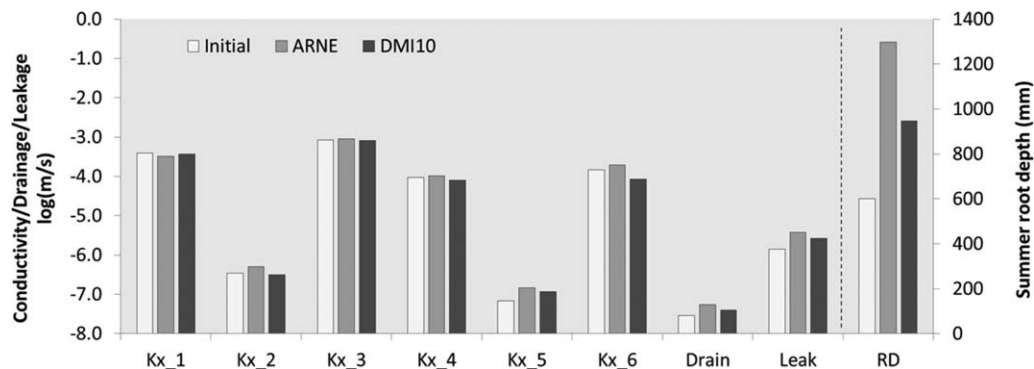


**Figure 8.** Calibrated parameter values using different precipitation inputs for the RF model. The explanations for the parameter symbols can be seen in Table 1. Initial parameter values are obtained from the calibration of a larger scale hydrological model using the same model code where the RF model is included. All values except for the root depth are log-transformed.

**Table 3.** Model Performance for Simulated Stream Discharge Using Different Precipitation Inputs During Both Calibration and Validated Periods, Expressed by the Nash-Sutcliffe Coefficient, and Deviations to Water Balance

| | Station | Area | NSE (–) | | $F_{bal\_tol}$ (%) | | $F_{bal\_jja}$ (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | ARNE | DMI10 | ARNE | DMI10 | ARNE | DMI10 |
| 2006 | 250021 | 46.5 | 0.79 | 0.68 | 15.02 | 21.73 | 13.77 | 24.81 |
| | 250086 | 80.6 | 0.70 | 0.73 | 4.96 | 2.22 | 2.92 | 2.51 |
| | 250018 | 81.9 | 0.23 | 0.22 | 16.68 | 17.01 | 22.06 | 21.67 |
| | 250075 | 91.8 | 0.58 | 0.55 | 5.76 | 6.28 | 26.35 | 25.33 |
| | 250020 | 117.3 | 0.65 | 0.67 | 3.67 | 0.58 | 1.14 | 0.29 |
| | 250078 | 611.7 | 0.79 | 0.80 | 5.36 | 2.87 | 5.07 | 2.32 |
| | 250082 | 1054.6 | 0.84 | 0.83 | 6.56 | 8.33 | 1.40 | 3.02 |
| | 250097 | 1550.0 | 0.74 | 0.74 | 8.15 | 8.80 | 4.48 | 4.69 |
| 2007–2009 | 250021 | 46.5 | 0.73 | 0.67 | 6.34 | 12.30 | 11.64 | 5.83 |
| | 250086 | 80.6 | 0.72 | 0.63 | 1.66 | 1.35 | 6.20 | 7.95 |
| | 250018 | 81.9 | 0.23 | 0.42 | 1.81 | 0.88 | 2.96 | 5.08 |
| | 250075 | 91.8 | 0.52 | 0.51 | 5.67 | 4.90 | 16.20 | 15.42 |
| | 250020 | 117.3 | 0.47 | 0.39 | 17.96 | 17.74 | 14.65 | 11.63 |
| | 250078 | 611.7 | 0.63 | 0.65 | 7.28 | 2.23 | 0.70 | 0.10 |
| | 250082 | 1054.6 | 0.76 | 0.72 | 4.41 | 2.73 | 2.19 | 1.87 |
| | 250097 | 1550.0 | 0.77 | 0.72 | 0.44 | 0.19 | 1.59 | 0.48 |
| 2010 | 250021 | 46.5 | 0.25 | 0.24 | 3.88 | 12.37 | 22.80 | 10.50 |
| | 250086 | 80.6 | 0.26 | 0.16 | 10.14 | 0.34 | 8.43 | 1.63 |
| | 250075 | 91.8 | 0.15 | 0.19 | 14.65 | 4.73 | 46.88 | 36.68 |
| | 250020 | 117.3 | −0.38 | −0.25 | 28.30 | 24.48 | 24.89 | 18.77 |
| | 250078 | 611.7 | 0.66 | 0.62 | 8.60 | 11.44 | 17.79 | 24.61 |
| | 250082 | 1054.6 | 0.69 | 0.62 | 0.27 | 3.33 | 1.00 | 4.11 |
| | 250097 | 1550.0 | 0.61 | 0.62 | 1.64 | 0.82 | 3.14 | 3.46 |

of the two-layer model ends at the root depth, it is likely that the buffering capacity is insufficient or the time delay in the unsaturated zone is too short. As a result, the infiltration excessive runoff is discharged directly to the rivers, whereas in real nature the catchment enables more buffering capacity than in the model.

[43] In the validation periods, 2006 and 2010, the differences between models performances are relatively small, which suggests that outside the calibration period the differences in predictive accuracy with respect to stream discharge caused by using different rainfall products are very limited. The difference between the two models at the same station is much smaller than the variation for the same

**Table 4.** Root Mean Squared Error of the Simulated Hydraulic Head at Different Model Computational Layers Using Radar and Rain Gauge Based Precipitation Inputs During the Calibration and Validation Periods (unit: m)[a]

| | No. of | 2006 | | 2007–2009 | | 2010 | |
|---|---|---|---|---|---|---|---|
| Layer | Observations | ARNE | DMI10 | ARNE | DMI10 | ARNE | DMI10 |
| 1 | 42 | 3.0 | 3.1 | 3.3 | 3.4 | 3.2 | 3.2 |
| 2 | 13 | 3.9 | 3.9 | 3.7 | 3.7 | 4.0 | 3.9 |
| 3 | 155 | 2.9 | 3.0 | 3.0 | 3.1 | 3.0 | 3.0 |
| 4 | 52 | 3.0 | 3.0 | 2.9 | 3.0 | 3.0 | 3.0 |
| 5 | 139 | 2.7 | 2.7 | 2.6 | 2.6 | 2.6 | 2.6 |
| 6 | 48 | 3.5 | 3.6 | 3.6 | 3.7 | 3.5 | 3.6 |
| 7 | 161 | 2.9 | 2.9 | 3.0 | 3.0 | 3.0 | 3.0 |
| 8 | 111 | 3.6 | 3.7 | 3.8 | 3.9 | 3.7 | 3.7 |
| 9 | 35 | 4.2 | 4.5 | 4.0 | 4.4 | 4.0 | 4.4 |
| 10 | 61 | 3.8 | 3.9 | 4.0 | 4.1 | 3.9 | 4.0 |
| 11 | 6 | 4.3 | 4.3 | 4.4 | 4.4 | 4.5 | 4.6 |

[a]Calculations are made with reference to the mean hydraulic head in 2000–2005.

model at different subcatchment. In addition, it was expected that the performance of the radar-based model would be better in smaller catchments due to the higher resolution. However, such scale dependency is unclear. One possible explanation is that the selected catchments are simply too large to verify this hypothesis. The smallest catchment covers an area of 46.5 km[2], which corresponds to nearly 12 radar pixels, or half of an interpolated rain gauge grid. Therefore, as long as the mean precipitation over this area is basically the same between the two estimates, the internal distribution does not make a significant impact.

### 4.3.2. Groundwater Head

[44] Using multiannual radar QPE data set as input to a hydrological model provides an opportunity to evaluate the impact of radar QPE on the simulated groundwater behavior, since the interaction between rainfall and groundwater is much less instantaneous than the rainfall-runoff process and thus requires much longer period of rainfall data set. Table 4 shows the RMSE of groundwater hydraulic head calculated for the models using both radar and rain-gauge-based precipitation inputs, and compared against observations from a control period (2000–2005) where abundant data are available. The model performances based on the two rainfall scenarios are very close to one another both in the calibration and validation periods with a general preference given to the ARNE-based model. It is noted that at layer 3, where the shallow aquifer is located, the performance of the radar-based model is about 10% higher than the rain-gauge-based model in all three periods. Since the observations from groundwater wells are relatively local, the advantage introduced by using high-resolution rainfall data as model input is apparent for simulating shallow groundwater heads. At deeper groundwater layers, for instance layer 5 and 8, the results from the radar-based

**Table 5.** Overall Water Balance for the Optimized Hydrological Models Over the Entire RF Catchment Using Different Precipitation Inputs During Both Calibration and Validation Periods (unit: mm/year)

| | 2006 | | 2007–2009 | | 2010 | |
|---|---|---|---|---|---|---|
| | ARNE | DMI10 | ARNE | DMI10 | ARNE | DMI10 |
| Precipitation | 1141 | 1134 | 1124 | 1111 | 915 | 914 |
| Evapotranspiration | 555 | 550 | 590 | 588 | 553 | 551 |
| Stream discharge | 411 | 413 | 503 | 498 | 368 | 370 |
| Base flow | 243 | 244 | 291 | 292 | 234 | 236 |
| Drain flow | 113 | 99 | 147 | 125 | 102 | 91 |
| Overland flow | 55 | 70 | 65 | 82 | 32 | 43 |
| Groundwater pumping | 18 | 23 | 28 | 36 | 22 | 27 |
| Groundwater storage change | 159 | 154 | −1 | −1 | −35 | −36 |
| Error | 9 | 15 | 23 | 31 | 16 | 20 |

model are also superior. However, although the groundwater recharge could dominate the simulated hydraulic head, it is a very indirect proxy of precipitation. Therefore, it is hardly conclusive if the improvement is caused by using radar QPE as the rainfall input.

[45] The radar and rain gauge simulated groundwater heads for 2006–2010 are illustrated in Figure 10. It is unnecessary to differentiate between the three periods since the differences are subtle. The groundwater head data are extracted from layer 3 of the saturated zone which is located approximately 10 to 20 m below the ground surface. On average, the spatial distribution of the simulated groundwater heads is very similar by visual comparison. However, the radar-driven hydrological model predicts higher groundwater elevation in the western part of the catchment and lower in the eastern part compared to the results obtained by the rain-gauge-driven model. Initially, it was anticipated that the choice of rainfall product would have limited effect on the simulated groundwater head over a time span of half a decade. However, Figure 10 indicates that the spatial distribution of precipitation applied in the hydrological model has relatively significant impact on the simulated groundwater heads: at many locations in the catchment the absolute differences between model simulations are more than 2 m. Since the RMSE of the simulated groundwater head is in the range of 2–4 m as indicated by Table 4, a 2 m difference between the model simulations could show significant impact on the model performances.
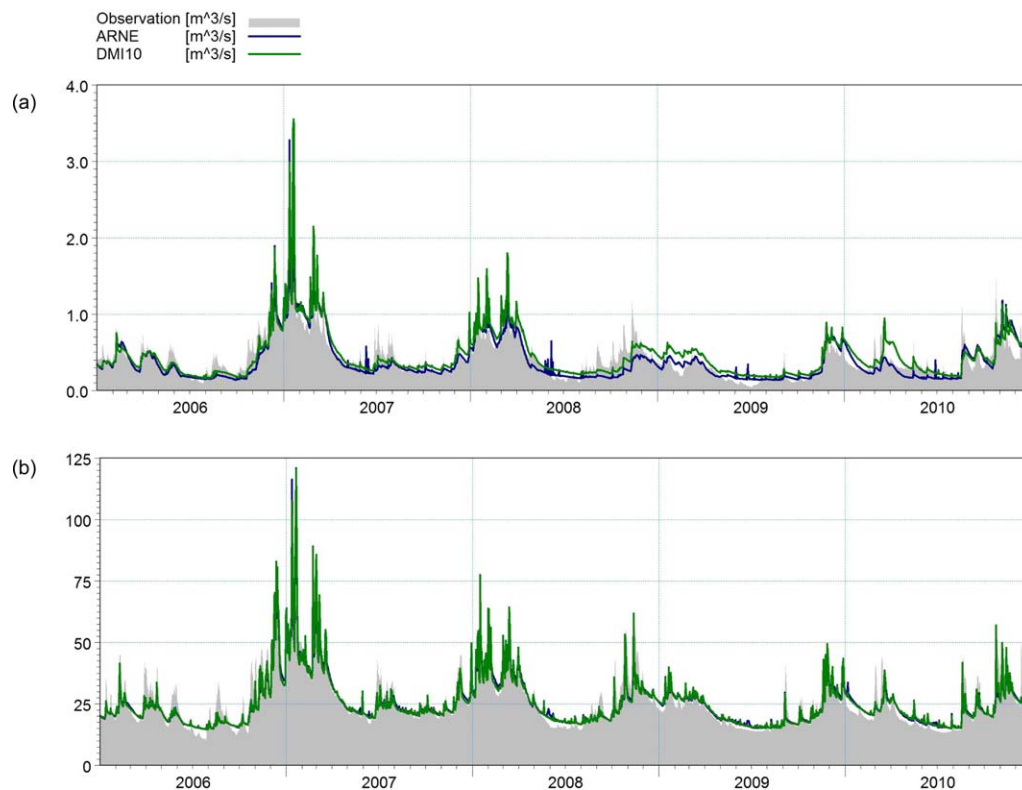


**Figure 9.** Hydrographs showing simulated stream discharge at (a) discharge station 250021 and (b) station 250097 at the Skjern River. The simulated discharge is obtained from the hydrological models calibrated individually using ARNE and DMI10.
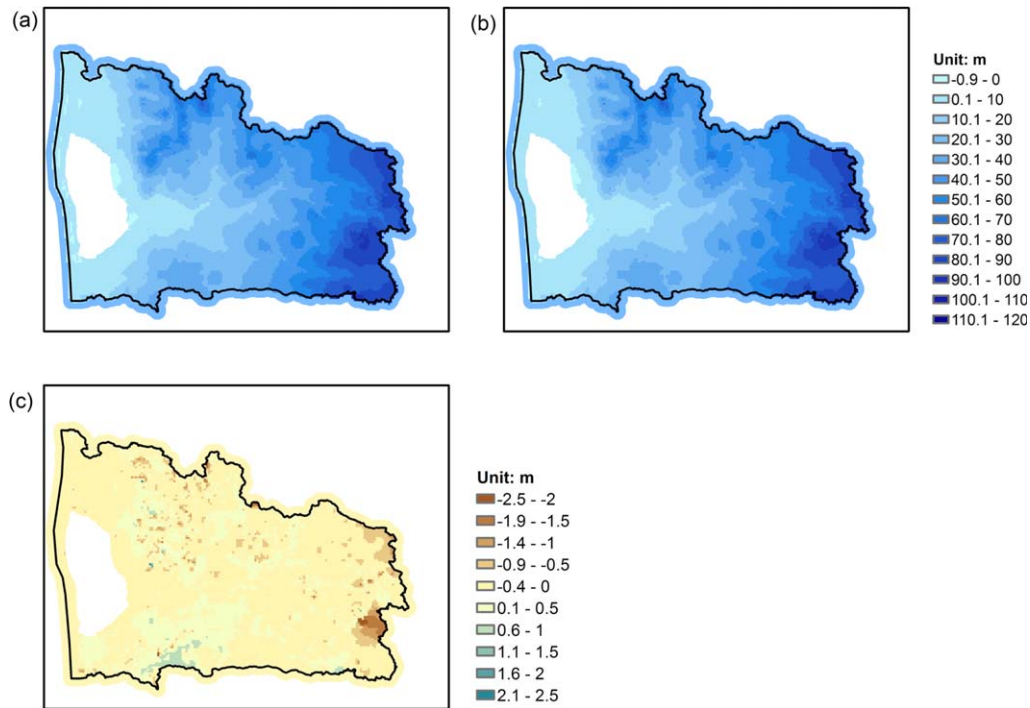
**Figure 10.** Maps showing simulated average groundwater head (2006–2010) extracted from layer 3 of the saturated zone by using (a) ARNE as precipitation input, (b) DMI10 as precipitation input to the hydrological model, and (c) difference between Figures 10a and 10b.

### 4.3.3. Overall Hydrological Model Performances

[46] Overall water balances for the two models are summarized in Table 5 for both the calibration and validation periods. The average precipitation estimated by ARNE is almost identical to DMI10 in our testing period with only marginal residual bias between the two products. As a result, Table 5 indicates that the two precipitation products have basically equal performance in providing sufficient accuracy for closing the water balance. It should also be noted that the model calibration has compensated for some of the differences in the rainfall forcing. Therefore, if an uncalibrated model with the same set of parameters is used for both precipitation estimates, the difference between the model outcomes would have been larger.

[47] In order to rate the model performance for different rain gauge layouts over time, measures for aggregation and averaging are needed. In the present study, RMSEs of the groundwater heads are calculated using time series observations, as seen in Figure 2a (2). RMSE scores from each groundwater well are weighted equally, whereas for the surface water weighted mean Nash-Sutcliffe efficiencies are used. The weighting is done considering the size of the subcatchments represented by the discharge stations, and the weighting factors are directly proportional to the subcatchment size. The results are shown in Figure 11.

[48] Both the radar and the rain-gauge-based models indicate somehow dependency on the rain gauge density with increasing RMSEs for groundwater head and decreasing NSE for stream discharge with decreasing density. RMSE increases by 16 cm and 28 cm from 2006 to 2010 for ARNE and DMI10, respectively. For the Nash-Sutcliffe coefficient, a decrease of 0.17 and 0.20 for the two data

sets are found. Hence, the decrease in performance is more apparent for the rain-gauge-based model.

[49] Relative changes in the performance criteria RMSE for groundwater and NSE for stream discharge are also apparent for simulations with and without the use of radar data. In the calibration period (2007–2009), both groundwater head and stream discharge are better described when radar QPE is used as input to the hydrological model, where a decrease in RMSE of 6% and an increase in NSE by 4.4% are found. In 2006, where the number of gauges is relatively high, the added value of using radar data is relatively small, 3.9% for groundwater and none for stream discharge. For the last period, 2010, where the gauge density is relatively low, the opposite result is found as the increase in performance of both RMSE and NSE are higher than for the calibration period. Hence, the results indicate that the value of the radar data increase as the gauge density decrease.

### 4.4. Model With Reduced Rain Gauge Density in 2006

[50] The weak model performance observed in 2010 could be due to abrupt changes in environmental condition in comparison to the previous period. If so, the decline in model performance cannot be explained by the change in rain gauge density. Therefore, the ARNE algorithm is run for the year 2006 where only rain gauges that are still in operation in 2010 are included. The results are shown by the red bars in Figure 11. The drop in rain gauge number has a clear negative influence on both the simulated groundwater head and stream discharge performances compared to the 2010 situation. The relative change in the performance criteria RMSE and NSE are calculated by
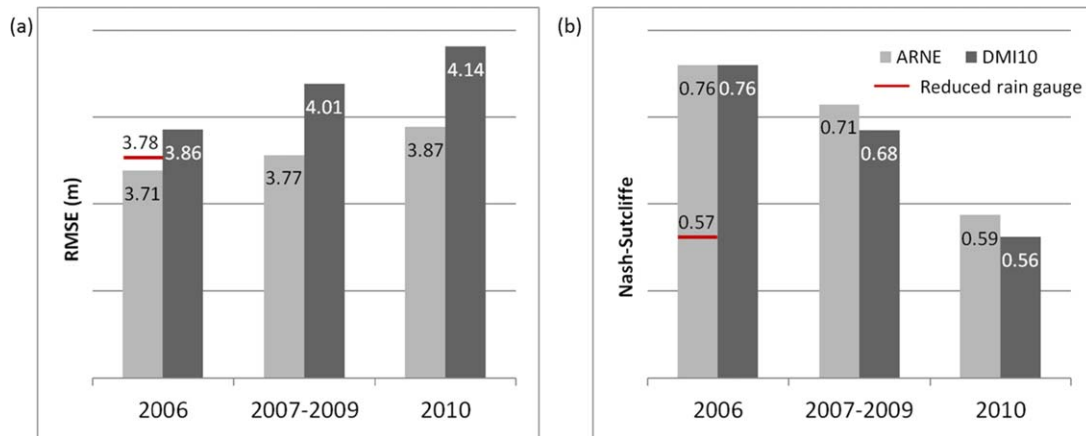
**Figure 11.** Comparisons of model performance in calibration and validation periods using radar and rain-gauge-based precipitation products as model inputs: (a) RMSE of simulated groundwater head and (b) Weighted mean Nash-Sutcliffe coefficients of simulated stream discharge.

comparing the results of using the full gauge network from 2006 to the reduced network from 2006 and the actual network from 2010. In both situations, substantial reductions in stream discharge performance of 25% and 22% are found, whereas the effect on groundwater is less significant, only an increase of 1.9% and 4.3% are found in RMSE. Therefore, reducing the rain gauge density in the QPE algorithm has larger impact on simulated stream discharge than that on simulated groundwater head.

[51] The change of model performance caused by the change of rain gauge density involved in the QPE algorithm also has spatial implications. The spatial arrangement of the rain gauges that are still in use in 2010 and the rain gauges that are removed between 2007 and 2009 are shown in Figure 12. It is seen that the closed stations are centered on two clusters: one to the south and another to the east of the RF catchment. The cluster to the south is not supposed to have large influence on the simulated discharge since the rainfall collected from that area can hardly be manifested by either of the discharge stations. However, the east cluster is very close to the upstream discharge stations, where two of the stations, namely 250018 and 250021, show considerable decrease of NSE as shown in Table 6. Therefore, when the rain gauge number is halved in 2006, the impact is more significant on the simulated upstream discharge in the RF catchment.

[52] The large reduction of NSE seen from the upstream stations is likely due to a drastic change of estimated rainfall mean. This hypothesis is confirmed by comparing the actual rain gauge measurements with the ARNE estimated rainfall with two sets of rain gauge densities at five locations shown in Figure 12 and Table 7. The correlation coefficients between the two rainfall scenarios are nearly identical at all five stations indicating the temporal dynamics of ARNE are unchanged regardless of the number of rain gauges involved. However, very large differences are observed on the mean bias where ARNE with reduced rain gauge network has all negative tendencies. Therefore, Table 7 provides evidence that by removing half of the rain gauges from the QPE algorithm, the estimated mean rainfall can change considerably, leading to a performance

decrease of the hydrological model especially at streams located at the upstream part of the catchment.

## 5. Discussion and Conclusions

[53] The present study highlights the comparison of model simulations from an overall water balance perspective with much consideration given to the simulated groundwater behavior, which is made feasible by using a five year radar QPE data set. Furthermore, the high dependence of model performance on the areal mean precipitation addresses the importance of rain gauge density involved in the radar QPE algorithm.

[54] Radar QPE has been compared with rain-gauge-based rainfall estimation and also quality checked using a hydrological model. Based on accumulated rainfall maps, the differences between ARNE and DMI10 are small both with respect to mean bias and spatial patterns. However, considering the improvement in the spatial resolution, ARNE is more favorable than the DMI10 product from a hydrological modeling perspective since some of the local rainfall missed by the interpolation of the rain gauge data may be captured by the radar QPE.

[55] The lumped performance measures suggest overall better groundwater head simulations for the radar-based model than the rain-gauge-based model regardless of the number of rain gauges involved, where a difference of approximately 5% in RMSE can be expected between the two models. For simulation of stream discharge, the benefit outlined above may not have large influence in medium to large size river basins, such as the Skjern catchment and its subcatchments, mainly due to the spatial averaging and smoothing in the processes. The positive impact of radar-based rainfall estimation on groundwater simulation is likely explained by the relatively slow response from rainfall on groundwater head. The catchment also acts as a filter that screens part of the noises and artifacts in the radar QPE. Therefore, the simulated groundwater head is much less sensitive to the radar artifacts and clutters. Besides, the groundwater head measurements are more local than the discharge measurements. As a result, the rainfall spatial
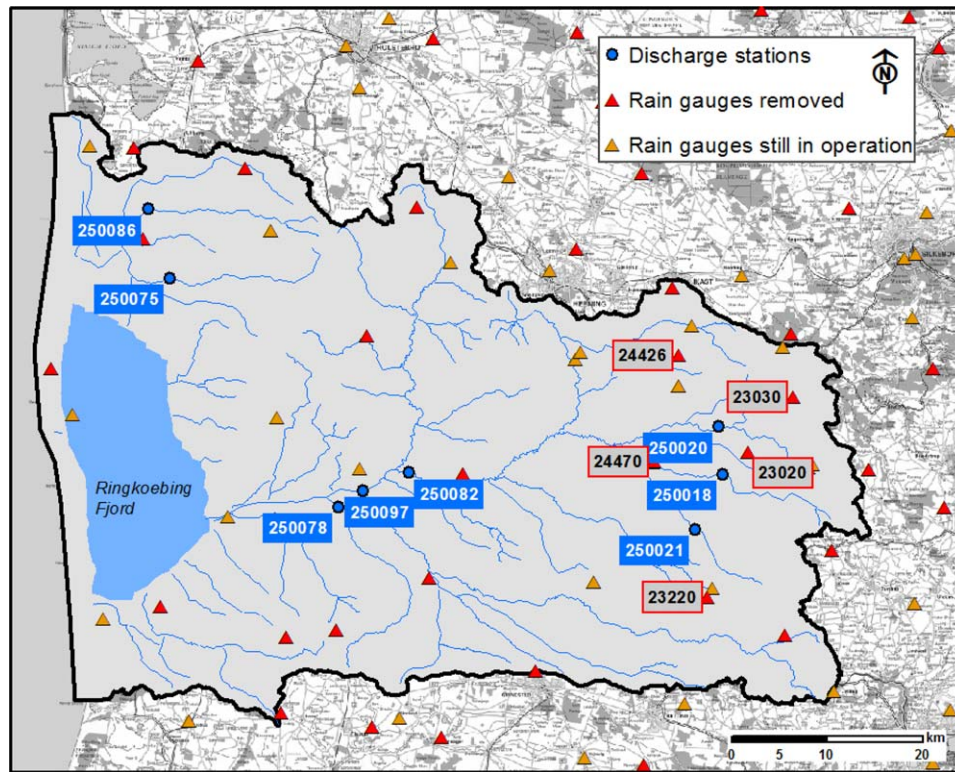
**Figure 12.** Illustration of rain gauges that are still in operation in 2010 (in yellow) and the rain gauges that have been removed between 2007–2009 (in red), as well as their relative locations to the stream discharge stations. Five rain gauges are chosen (station numbers are marked in red boxes) for further analysis where the results are seen in Table 7.

details revealed by using the radar data have demonstrated advantages over the rain-gauge-based precipitation estimates in simulation of groundwater heads.

[56] In hydrological studies, weather radar has been reportedly used to assist flood forecasting, since the infiltration excessive runoff is directly related to the intensity and spatial distribution of rainfall [*Kuczera and Williams*, 1992; *Sun et al.*, 2000; *Carpenter and Georgakakos*, 2004; *Ntelekos et al.*, 2006; *Cole and Moore*, 2008]. The present study, from a quantitative perspective, demonstrates the potential of using radar QPE for groundwater modeling at places where surface and subsurface hydrological processes are closely linked. This finding could help to extend the application of radar-

**Table 6.** Changes of Nash-Sutcliffe Coefficients of Simulated Stream Discharge Caused by Changes in Rain Gauge Density in Radar Based QPE in 2006

| Station | Area (km$^2$) | Full Network | Reduced Network | ΔNSE |
|---------|---------------|--------------|-----------------|------|
| 250021 | 47 | 0.79 | 0.33 | 0.47 |
| 250086 | 81 | 0.70 | 0.67 | 0.02 |
| 250018 | 82 | 0.23 | −0.45 | 0.69 |
| 250075 | 92 | 0.58 | 0.42 | 0.16 |
| 250020 | 117 | 0.65 | 0.47 | 0.18 |
| 250078 | 612 | 0.79 | 0.73 | 0.06 |
| 250082 | 1055 | 0.84 | 0.60 | 0.24 |
| 250097 | 1550 | 0.74 | 0.52 | 0.22 |

based QPE data to a broader range of applications, which otherwise have not been given enough attention.

[57] It has been suggested in the present study that the added value of using radar increases when the number of rain gauges decreases. In 2006, where abundant gauges are available, the differences in RMSE and NSE between the ARNE and DMI10 based simulations are relatively small. As the number of gauges is reduced from 2006 to 2010 the added value of using radar becomes apparent. We may assume that with the help from radar it is possible to sustain the quality of the resulting rainfall product when the rain gauge number reduces. However, our results indicate that the quality of the radar-based product is lower than the original gauge-based product which is contrary to the intention.

[58] To further analyze this problem, a synthetic scenario was created to reduce the rain gauge number in 2006 to the same level as in 2010, which provides an opportunity to analyze the impact of the rain gauge density on both the radar-based rainfall estimation and on the hydrological simulations under exactly the same climatic and hydrological conditions. It is implied that maintaining a relatively high rain gauge density, e.g., 4 gauges/100 km$^2$, is crucial even with the aid from the radar. With many uncertainties in the radar-based QPE still unsolved, we consider that rain gauges, especially the ones close to the upstream part of the RF catchment, should not have been removed for the time being from a hydrological modeling perspective. The NSE score obtained by using ARNE with reduced rain

**Table 7.** Comparisons Between Rain Gauge Observed Rainfall and ARNE Estimated Rainfall in 2006, With Full Rain Gauge Network and Reduced Rain Gauge Network in ARNE

| Station | Mean Bias (mm/year) | | RMSE (mm/year) | | Correl. Coefficient (r) | |
| --- | --- | --- | --- | --- | --- | --- |
| | Full Network | Reduced Network | Full Network | Reduced Network | Full Network | Reduced Network |
| 23020 | 75.40 | −36.96 | 349.44 | 383.25 | 0.98 | 0.98 |
| 23030 | −56.03 | −89.25 | 345.06 | 402.55 | 0.98 | 0.98 |
| 23220 | 41.38 | −48.47 | 598.83 | 678.05 | 0.94 | 0.92 |
| 24426 | 33.69 | −65.60 | 347.42 | 376.57 | 0.98 | 0.98 |
| 24470 | 13.44 | −87.31 | 504.57 | 531.62 | 0.97 | 0.97 |

gauge network in 2006 from discharge station 250021 is similar to the NSE score from the same discharge station in 2010, which are 0.33 and 0.25, respectively. Therefore, the negative impact on the upstream station by closing down the neighborhood rain gauges is consistent. It can be concluded that closing down half of the rain gauge stations in the study area sacrifices the quality of the rainfall estimation and subsequently the performance of the hydrological model; however, without the extra data obtained from the radar the effect would have been worse.

[59] There are mainly two reasons that cause model performance to decrease when using fewer rain gauges in the QPE algorithm. First, it could be due to the change of estimated rainfall mean. Several studies have indicated that in distributed hydrological modeling, the estimated rainfall mean plays a significant role in sustaining realistic model simulations, especially when the model results have to be extrapolated [*Bandaragoda et al.*, 2004; *Ivanov et al.*, 2004; *Jayawickreme and Hyndman*, 2007]. Our study has confirmed this hypothesis that when the estimated mean precipitation changes from 1141 mm/yr to 1086 mm/yr in 2006, which is caused by the reduction of rain gauge density from around 4 gauges/100 km$^2$ to 2 gauges/100 km$^2$, there is also a significant decrease in model performance. Second, less accurate characterization of the spatial variation of rainfall (rainfall noise) could also make a negative impact on the hydrological simulation. Model calibration can usually to a large extent compensate for the rainfall bias, but is much less effective to the rainfall noise. With the model calibrated for 2007–2009, one could expect poorer performance when validating for a different time period. This is true in 2010. However, in 2006, more rainfall stations not only help to reduce the rainfall bias but also the rainfall noise, which results in a better performance than the calibration period. We believe that both of the two factors are important, and their combined effect has the final impact on the hydrological simulations.

[60] The QPE production method applied in the present study can also have implications on the conclusions. One of the main objectives in QPE is to make adjustments to the radar rainfall images and bring them closer to what can be observed at ground level. To achieve that, individual uncertainty sources are studied and physically based corrections are made [*Bech et al.*, 2000; *Vignal et al.*, 2000; *Anagnostou et al.*, 2004; *Brandes et al.*, 2004]. Rain gauge data are not always involved in this process and thus the reduction of rain gauge numbers will not pose any significant impact. An alternative approach is to consider all the uncertainty sources as a whole, and this usually requires

the use of rain gauge data. In the present study, the second approach is taken.

[61] The overall conclusions from the present study are expected to be valid at other locations with similar hydrological regimes, i.e., rather homogeneous precipitation, flat terrain, and groundwater dominated hydrology. The same trend in terms of rain gauge density would also be expected in mountainous areas or other areas with higher spatial variation in terrain and precipitation.

# References

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. Oconnell, and J. Rasmussen (1986), An introduction to the European hydrological system - Systeme Hydrologique Europeen, SHE. 1. History and philosophy of a physically-based, distributed modeling system, *J. Hydrol.*, *87*(1-2), 45–59.

Abtew, W., J. Obeysekera, and G. Shih (1995), Spatial variation of daily rainfall and network design, *T Asae*, *38*(3), 843–845.

Anagnostou, E. N., and W. F. Krajewski (1998), Calibration of the wsr-88d precipitation processing subsystem, *Weather Forecasting*, *13*(2), 396–406.

Anagnostou, E. N., W. F. Krajewski, D. J. Seo, and E. R. Johnson (1998), Mean-field rainfall bias studies for wsr-88d, *J. Hydrol. Eng.*, *3*(3), 149–159.

Anagnostou, E. N., M. N. Anagnostou, W. F. Krajewski, A. Kruger, and B. J. Miriovsky (2004), High-resolution rainfall estimation from x-band polarimetric radar measurements, *J. Hydrometeorol.*, *5*(1), 110–128.

Atencia, A., L. Mediero, M. C. Llasat, and L. Garrote (2011), Effect of radar rainfall time resolution on the predictive capability of a distributed hydrologic model, *Hydrol. Earth Syst. Sci.*, *15*(12), 3809–3827.

Bandaragoda, C., D. G. Tarboton, and R. Woods (2004), Application of topnet in the distributed model intercomparison project, *J. Hydrol.*, *298*(1–4), 178–201.

Barnes, S. L. (1964), A technique in maximizing details in numerical weather map analysis, *J. Appl. Meteorol.*, *3*(4), 396–409.

Bech, J., A. Sairouni, B. Codina, J. Lorente, and D. Bebbington (2000), Weather radar anaprop conditions at a Mediterranean coastal site, *Phys. Chem. Earth Part B*, *25*(10-12), 829–832.

Benito, G., et al. (2010), Management of alluvial aquifers in two Southern African ephemeral rivers: Implications for IWRM, *Water Resour. Manage*, *24*(4), 641–667.

Beven, K. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrol. Processes*, *16*, 189–206.

Borga, M., and F. Tonelli (2000), Adjustment of range-dependent bias in radar rainfall estimates, *Phys. Chem. Earth Part B*, *25*(10–12), 909–914.

Borga, M., F. Tonelli, R. J. Moore, and H. Andrieu (2002), Long-term assessment of bias adjustment in radar rainfall estimation, *Water Resour. Res.*, *38*(11), 1226, doi:10.1029/2001WR000555.

Bormann, H., B. Diekkruger, and C. Renschler (1999), Regionalisation concept for hydrological modelling on different scales using a physically based model: Results and evaluation, *Phys. Chem. Earth Part B*, *24*(7), 799–804.

Boyle, D. P., H. V. Gupta, S. Sorooshian, V. Koren, Z. Y. Zhang, and M. Smith (2001), Toward improved streamflow forecasts: Value of semidistributed modeling, *Water Resour. Res.*, *37*(11), 2749–2759.

Brandes, E. A. (1975), Optimizing rainfall estimates with aid of radar, *J. Appl. Meteorol.*, *14*(7), 1339–1345.

Brandes, E. A., G. F. Zhang, and J. Vivekanandan (2004), Drop size distribution retrieval with polarimetric radar: Model and application, *J. Appl. Meteorol.*, *43*(3), 461–475.

Burns, D. A., J. J. McDonnell, R. P. Hooper, N. E. Peters, J. E. Freer, C. Kendall, and K. Beven (2001), Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola mountain research watershed (Georgia, USA), *Hydrol Processes*, *15*(10), 1903–1924.

Carpenter, T. M., and K. P. Georgakakos (2004), Continuous streamflow simulation with the hrcdhm distributed hydrologic model, *J. Hydrol.*, *298*(1–4), 61–79.

Carpenter, T. M., J. A. Sperfslage, K. P. Georgakakos, T. Sweeney, and D. L. Fread (1999), National threshold runoff estimation utilizing gis in support of operational flash flood warning systems, *J. Hydrol.*, *224*(1–2), 21–44.

Christensen, S., and J. Doherty (2008), Predictive error dependencies when using pilot points and singular value decomposition in groundwater model calibration, *Adv. Water Resour.*, *31*(4), 674–700.

Ciach, G. J., W. F. Krajewski, E. N. Anagnostou, M. L. Baeck, J. A. Smith, J. R. McCollum, and A. Kruger (1997), Radar rainfall estimation for ground validation studies of the tropical rainfall measuring mission, *J. Appl. Meteorol.*, *36*(6), 735–747.

Clark, R. A., Y. J. Canipe, and D. R. Greene (1972), Applications of digital radar data in both meteorology and hydrology, *B. Am. Meteorol. Soc.*, *53*(7), 687.

Cole, S. J., and R. J. Moore (2008), Hydrological modelling using raingauge- and radar-based estimators of areal rainfall, *J. Hydrol.*, *358*(3–4), 159–181.

Cole, S. J., and R. J. Moore (2009), Distributed hydrological modelling using weather radar in gauged and ungauged basins, *Adv. Water Resour.*, *32*(7), 1107–1120.

Doherty, J. (2003), Ground water model calibration using pilot points and regularization, *Ground Water*, *41*(2), 170–177.

Fu, S. H., T. O. Sonnenborg, K. H. Jensen, and X. He (2011), Impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resources model, *Vadose Zone J.*, *10*(1), 25–36.

Goudenhoofdt, E., and L. Delobbe (2009), Evaluation of radar-gauge merging methods for quantitative precipitation estimates, *Hydrol. Earth Syst. Sci.*, *13*(2), 195–203.

Gourley, J. J., and B. E. Vieux (2005), A method for evaluating the accuracy of quantitative precipitation estimates from a hydrologic modeling perspective, *J. Hydrometeorol.*, *6*(2), 115–133.

Gourley, J. J., S. E. Giangrande, Y. Hong, Z. L. Flamig, T. Schuur, and J. A. Vrugt (2010), Impacts of polarimetric radar observations on hydrologic simulation, *J. Hydrometeorol.*, *11*(3), 781–796.

Graham, D. N., and M. B. Butts (2006), Flexible, integrated watershed modelling with MIKE SHE, in *Watershed Models*, edited by V. P. Singh and D. K. Frevert, pp. 245–272, Taylor and Francis, London, United Kingdom.

Gupta, H. V., L. A. Bastidas, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang (1999), Parameter estimation of a land surface scheme using multicriteria methods, *J. Geophys. Res.*, *104*(D16), 19,491–19,503.

Haberlandt, U. (2007), Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, *J. Hydrol.*, *332*(1–2), 144–157.

Haberlandt, U., B. Klocking, V. Krysanova, and A. Becker (2001), Regionalisation of the base flow index from dynamically simulated flow components—A case study in the elbe river basin, *J. Hydrol.*, *248*(1-4), 35–53.

He, X., F. Vejen, S. Stisen, T. O. Sonnenborg, and K. H. Jensen (2011a), An operational weather radar-based quantitative precipitation estimation and its application in catchment water resources modeling, *Vadose Zone J.*, *10*(1), 8–24.

He, X., J. C. Refsgaard, T. O. Sonnenborg, F. Vejen, and K. H. Jensen (2011b), Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling, *Water Resour Res*, *47*, W09526, doi:10.1029/2011WR010670.

Henriksen, H. J., L. Troldborg, P. Nyegaard, T. O. Sonnenborg, J. C. Refsgaard, and B. Madsen (2003), Methodology for construction, calibration and validation of a national hydrological model for Denmark, *J. Hydrol.*, *280*(1–4), 52–71.

Henriksen, H. J., L. Troldborg, A. L. Hojberg, and J. C. Refsgaard (2008), Assessment of exploitable groundwater resources of denmark by use of ensemble resource indicators and a numerical groundwater-surface water model, *J. Hydrol.*, *348*(1–2), 224–240.

Hojberg, A. L., L. Troldborg, S. Stisen, B. S. B. Christensen, and H. J. Henriksen (2012), Stakeholder driven update and improvement of a national water resources model, *Environ. Modell. Software*, *40*, 202–213.

Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi (2004), Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: A fully-distributed physically-based approach, *J. Hydrol.*, *298*(1–4), 80–111.

Jayawickreme, D. H., and D. W. Hyndman (2007), Evaluating the influence of land cover on seasonal water budgets using next generation radar (nexrad) rainfall and streamflow data, *Water Resour Res*, *43*, W02408, doi:10.1029/2005WR004460.

Jensen, K. H., and T. H. Illangasekare (2011), Hobe: A hydrological observatory, *Vadose Zone J.*, *10*(1), 1–7.

Kim, K., K. Kim, and C. Yoo (2010), Decision of minimum rain gauge density in a combined radar-rain gauge rainfall observation system: A case study of the Imjin river basin, Korea, *Water Int.*, *35*(1), 49–62.

Kitchen, M., R. Brown, and A. G. Davies (1994), Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation, *Q. J. R. Meteorol. Soc.*, *120*(519), 1231–1254.

Krajewski, W. F., and B. Vignal (2001), Evaluation of anomalous propagation echo detection in WSR-88D data: A large sample case study, *J. Atmos. Ocean Tech.*, *18*(5), 807–814.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S. C. Jain (1991), A Monte-Carlo study of rainfall sampling effect on a distributed catchment model, *Water Resour. Res.*, *27*(1), 119–128.

Krause, S., A. Bronstert, and E. Zehe (2007), Groundwater-surface water interactions in a north german lowland floodplain—Implications for the river discharge dynamics and riparian water balance, *J. Hydrol.*, *347*(3–4), 404–417.

Kuczera, G., and B. J. Williams (1992), Effect of rainfall errors on accuracy of design flood estimates, *Water Resour. Res.*, *28*(4), 1145–1153.

Liang, X., and Z. H. Xie (2003), Important factors in land-atmosphere interactions: Surface runoff generations and interactions between surface and groundwater, *Global Planet. Change*, *38*(1–2), 101–114.

Ludwig, R. et al. (2003), Web-based modelling of energy, water and matter fluxes to support decision making in mesoscale catchments—The integrative perspective of Glowa-Danube, *Phys. Chem. Earth*, *28*(14–15), 621–634.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Adv. Water Resour.*, *26*(2), 205–216.

Marshall, J. S., R. C. Langille, W. M. Palmer, and D. W. Perrie (1947), The measurement of rainfall by radar, *Phys. Rev.*, *72*(6), 538–538.

Michelson, D. B., and J. Koistinen (2000), Gauge-radar network adjustment for the baltic sea experiment, *Phys. Chem. Earth Part B*, *25*(10–12), 915–920.

Milewski, A., M. Sultan, E. Yan, R. Becker, A. Abdeldayem, F. Soliman, and K. A. Gelil (2009), A remote sensing solution for estimating runoff and recharge in arid environments, *J. Hydrol.*, *373*(1–2), 1–14.

Moore, C., and J. Doherty (2005), Role of the calibration process in reducing model predictive error, *Water Resour. Res.*, *41*, W05020, doi:10.1029/2004WR003501.

Ntelekos, A. A., K. P. Georgakakos, and W. F. Krajewski (2006), On the uncertainties of flash flood guidance: Toward probabilistic forecasting of flash floods, *J. Hydrometeorol.*, *7*(5), 896–915.

Nyegaard, P., L. Troldborg, and A. L. Højberg (2010), Geologisk og hydrostratigrafisk opdatering 2005–2009 (in Danish), GEUS Rep. Geological Survey of Denmark and Greenland, Copenhagen, Denmark [Available at http://vandmodel.dk/xpdf/dk-model2009_geol-hydr-opdatering.pdf.]

Refsgaard, J. C., and B. Storm (1995), MIKE SHE, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 804–846, Water Resour. Publ., Colo.

Ryzhkov, A., and D. S. Zrnic (1995), Precipitation and attenuation measurements at a 10-cm wavelength, *J. Appl. Meteorol.*, *34*(10), 2121–2134.

Scharling, M., K. Rajakumar, L. Hansen, and J. J. Jensen (2006), Cataloque of meteorological observing stations operated by DMI, DMI Internal Rep., Danish Meteorological Institute, Copenhagen, Denmark [Available at: http://www.dmi.dk/dmi/tr06–11.pdf.].

Seifert, D., T. O. Sonnenborg, J. C. Refsgaard, A. L. Hojberg, and L. Troldborg (2012), Assessment of hydrological model predictive ability given

multiple conceptual geological models, *Water Resour. Res.*, *48*, W06503, doi:10.1029/2011WR011149.

Seo, D. J., and J. P. Breidenbach (2002), Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements, *J. Hydrometeorol.*, *3*(2), 93–111.

Seo, D. J., J. P. Breidenbach, and E. R. Johnson (1999), Real-time estimation of mean field bias in radar rainfall data, *J. Hydrol.*, *223*(3–4), 131–147.

Skahill, B. E., and J. Doherty (2006), An advanced regularization methodology for use in watershed model calibration, *J. Hydrol.*, *327*(3–4), 564–577.

Smith, J. A., and W. F. Krajewski (1991), Estimation of the mean field bias of radar rainfall estimates, *J. Appl. Meteorol.*, *30*(4), 397–412.

Steiner, M., R. A. Houze, and S. E. Yuter (1995), Climatological characterization of 3-dimensional storm structure from operational radar and rain-gauge data, *J. Appl. Meteorol.*, *34*(9), 1978–2007.

Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden (1999), Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation, *Water Resour. Res.*, *35*(8), 2487–2503.

Stisen, S., T. O. Sonnenborg, A. L. Hojberg, L. Troldborg, and J. C. Refsgaard (2011), Evaluation of climate input biases and water balance issues using a coupled surface-subsurface model, *Vadose Zone J.*, *10*(1), 37–53.

Stisen, S., A. L. Højberg, L. Troldborg, J. C. Refsgaard, B. C. B. Christensen, M. Olsen, and H. J. Henriksen (2012), On the importance of appropriate rain-gauge catch correction for hydrological modelling at mid to high latitudes, *Hydrol. Earth Syst. Sci.*, *9*, 3607–3655.

Sun, X., R. G. Mein, T. D. Keenan, and J. F. Elliott (2000), Flood estimation using radar and raingauge data, *J. Hydrol.*, *239*(1-4), 4–18.

Tilford, K. A., N. I. Fox, and C. G. Collier (2002), Application of weather radar data for urban hydrology, *Meteorol. Appl.*, *9*(1), 95–104.

van de Beek, C. Z., H. Leijnse, J. N. M. Stricker, R. Uijlenhoet, and H. W. J. Russchenberg (2010), Performance of high-resolution x-band radar for rainfall measurement in the Netherlands, *Hydrol. Earth Syst. Sci.*, *14*(2), 205–221.

van Roosmalen, L., T. O. Sonnenborg, and K. H. Jensen (2009), Impact of climate and land use change on the hydrology of a large-scale agricultural catchment, *Water Resour. Res.*, *45*, W00A15, doi:10.1029/2007WR006760.

Vignal, B., G. Galli, J. Joss, and U. Germann (2000), Three methods to determine profiles of reflectivity from volumetric radar data to correct precipitation estimates, *J. Appl. Meteorol.*, *39*(10), 1715–1726.

Villarini, G., and W. F. Krajewski (2010), Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall, *Surv. Geophys.*, *31*(1), 107–129.

Villarini, G., F. Serinaldi, and W. F. Krajewski (2008), Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches, *Adv. Water Resour.*, *31*(12), 1674–1686.

Villarini, G., W. F. Krajewski, G. J. Ciach, and D. L. Zimmerman (2009), Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates, *Water Resour. Res.*, *45*, W01404, doi:10.1029/2008WR006946.

Watts, L. G., and A. Calver (1991), Effects of spatially-distributed rainfall on runoff for a conceptual catchment, *Nordic Hydrol.*, *22*(1), 1–14.

Wilson, J. W., and E. A. Brandes (1979), Radar measurement of rainfall—Summary, *Bull. Am. Meteorol. Soc.*, *60*(9), 1048–1058.

Younger, P. M., J. E. Freer, and K. J. Beven (2009), Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework, *Hydrol. Processes*, *23*(14), 1988–2003.

Zawadzki, I. (1984), Factors affecting the precision of radar measurement of rain, paper presented at 22nd International Conference on Radar Meteorology, *Am. Meteorol. Soc.*, Boston, Massachusetts.

Zoccatelli, D., M. Borga, A. Viglione, G. B. Chirico, and G. Bloschl (2011), Spatial moments of catchment rainfall: Rainfall spatial organisation, basin morphology, and flood response, *Hydrol. Earth Syst. Sci.*, *15*(12), 3767–3783.