PhD thesis

Xin He

Weather radar based quantitative precipitation estimation in modeling of catchment hydrology

Academic advisor: Prof. Karsten Høgh Jensen

31/July/2011
Preface

The thesis is a part of the requirements in order to complete the PhD degree at the Department of Geography and Geology (DGG), University of Copenhagen, Denmark. The PhD project is carried out in collaboration with the Danish Meteorological Institute (DMI) and the Geological survey of Denmark and Greenland (GEUS).

The work was carried out first at DMI for eighteen months and later at GEUS for another eighteen months. An external research stay of three month was spent at the National Center for Atmospheric Research, Boulder, Colorado. The PhD study has been supervised by Professor Karsten Høgh Jensen (DGG), Senior Scientist Flemming Vejen (DMI), Senior Scientist Torben Obel Sonnenborg (GEUS), and Professor Jens Christian Refsgaard (GEUS).

The PhD research and related activities have been a part of the HOBE project which was funded by the Villum Foundation (formerly Villum Kann Rasmussen Foundation).

Copenhagen, 2011

Xin He
Acknowledgements

It has been very fortunate for me to carry out this study in many interesting yet different research environments. Thus, there is a long list of people I would like to show my appreciation to.

First and foremost, I want to give my deepest gratitude to my principal supervisor Karsten Høgh Jensen for his unconditional support all along the way. It was not easy to be the first one trying to unlock the potential of the C-band weather radar for hydrological research in Denmark. His encouragement and leadership helped me to overcome all the obstacles. Flemming Vejen is also thanked for sharing his abundant knowledge on weather radar, as radar was an entirely new subject to me in the beginning.

I wish to give my special thank to Torben Obel Sonnenborg for his patience and support. Without him, none of the articles would have come to life. I am also very grateful to Jens Christian Refsgaard. He not only helped me to produce good science, but most importantly he inspired me to be a better scientist. Simon Stisen is also thanked for helping me with the hydrological model and the model optimizations, which seemed overwhelming at first glance.

I want to thank my colleagues at DMI, TK and the ‘radar group’ for making my stay there a pleasant experience. I also want to thank my colleagues at GEUS, especially my former/current officemates Peter Scharling, Britt Christensen, Martin Olsen for their great company, and Head of Department Heidi Christensen Barlebo for making the Department of Hydrology such a delightful place.

The motivating encouragements from Professor Tai-Chi Chen from National Central University are deeply appreciated. Shao-Fan Chang, Wei-Yu Chang from NCU and Shao-Chin Huang from Chinese Cultural University are also thanked for their warm hospitality both in Boulder and in Taiwan.

I wish to also thank my fellow PhD students and Assistant Professor Majken Looms Zibar at DGG for keeping me balanced between work and pleasure.

Last but not least, my families in China are thanked for their caring and moral support, and my friends in both China and Denmark are thanked for their faith and practical assistance.
Abstract

Ground based weather surveillance radars are able to provide quantitative precipitation estimation (QPE) that has high spatial and temporal accuracy. Using QPE data for hydrological applications has large potentials for improving the understanding of simulated hydrological processes in response to changes in precipitation forcing. The main objective of this PhD study is to combine the two scientific fields of radar QPE and catchment scale hydrological modeling. Previously, this kind of work has mostly been done by using event based radar data, and surface water has commonly been the hydrological component of interest. In the present PhD study, the long term effects of using continuous radar QPE data in a water resources modeling framework that includes both surface and subsurface hydrological components are analyzed.

The research activities have been carried out as a part of the HOBE project which is launched in the Skjern river catchment in western Denmark. Radar reflectivity data were obtained from three C-band weather radars operated by the Danish Meteorological Institute. A methodology was used to generate radar QPE images on a daily basis. The hydrological models were developed in line with the Danish Water Resource Model using the MIKE SHE model code.

The results confirm that the radar QPE data can be used as a reliable data source to serve as model inputs to distributed integrated hydrological models. The simulated hydrological responses by using the QPE data are comparable to the models driven by rain gauge data. The hydrological models show high sensitivity to the spatial distribution of rainfall in relatively small subcatchments, and the uncertainties in the radar rainfall show larger impact on the simulated stream discharge at small catchment scales. Although the QPE algorithm is able to produce rainfall estimates with relatively high quality, it is indicated in this study that in order to achieve satisfactory model performance, the radar data quality needs to be further improved.
Danske resumé

Data fra jordbaserede vejrradarer muliggør kvantitative estimator for nedbør (Quantitative Precipitation Estimation - QPE), der har en høj rumlig og tidlig opløsning. Radar QPE data har store anvendelsesmuligheder i hydrologisk sammenhæng, da en bedre regional bestemmelse af nedbørsinputtet til hydrologiske modeller vil kunne danne grundlag for en bedre kalibrering og beskrivelse af de hydrologiske processer samt af vandbalancen for et vandløbsopland. I modsætning til tidligere studier, hvor radar QPE primært har været anvendt i forbindelse med nedbørshændelser og de heraf følgende konsekvenser for vandløbsafstrømning, har nærværende PhD studium fokuseret på anvendelse af lange kontinuerede tidsserier for QPE. Disse tidsserier er anvendt som input til en distribueret hydrologisk model baseret på MIKE SHE koden, og det er undersøgt, hvorledes disse data påvirker simuleringen af afstrømning, fordampning, grundvandsdannelse og trykniveauer i grundvandet. Data er indsamlet fra tre C-bånd radarer, som drives af Danmarks Meteorologiske Institut (DMI). Radar QPE er genereret på daglig basis, og fokusområdet har været oplandet til Skjern å. Forskningen indgår som et element i HOBE projektet – et hydrologisk observatorium.

Resultaterne har vist, at radar QPE data kan anvendes som pålidelige input til distribuerede og integrerede hydrologiske modeller. De simulerede hydrologiske respons baseret på radar QPE data er sammenlignelige med de resultater, som opnås ved hjælp af data fra nedbørmålerne. For små deloplande udviser simuleringsresultaterne stor følsomhed for den rumlige fordeling af nedbøren. Ligeledes har usikkerheden på nedbørsestimaterne størst indflydelse på simuleringen af vandløbsafstrømningen for små deloplande. Selv om den anvendte radar QPE algoritme er i stand til at producere nedbørsestimater af relativt høj nøjagtighed, er der behov for at forbedre kvaliteten af radardata for dermed også at forbedre pålideligheden af de hydrologiske simuleringsresultater.
Contents

Preface ii
Acknowledgements iii
Abstract iv
Danske resumé iv

1. Motivations and objectives ................................................................. 1
2. Background ...................................................................................... 3
  2.1 Weather radar observation of the atmosphere .................................. 3
  2.2 Estimation of precipitation using weather radar ............................... 5
  2.3 Uncertainties in radar precipitation estimation ................................. 6
  2.4 Merging of radar and rain gauge data ............................................. 9
  2.5 Application of radar meteorology in hydrology .............................. 10
3. Introduction to thesis papers ............................................................. 12
  3.1 Introduction to Paper I ................................................................. 12
  3.2 Introduction to Paper II ............................................................... 13
  3.3 Introduction to Paper III .............................................................. 13
  3.4 Introduction to Paper IV .............................................................. 14
4. Conclusions and perspectives ......................................................... 15
References .......................................................................................... 17

Appendices

Paper I: An operational weather radar based QPE and its application in catchment water resources modeling

Paper II: Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling

Paper III: Calibration and validation of a distributed water resources model using radar and rain gauge based precipitation input

Paper IV: Impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resources model
1. Motivations and objectives

Precipitation is an essential part of the hydrological cycle in that it is the driving force to the other components of the overall water balance. In Denmark and other countries measurement of precipitation is routinely carried out by ground based rain gauge stations. As a result, the point based data sometimes fails to account for the spatial variability of the precipitation systems, which may cause inaccuracies in the hydrological modeling. Quantitative precipitation estimation (QPE) generated using both radar and rain gauge data have an ideal spatial resolution and high sampling frequency. Therefore, using such data may potentially improve the understanding and prediction of the hydrological processes.

As a work package within the Danish HOBE project (Jensen and Illangasekare 2011), the PhD study was offered a unique platform for generating precipitation products using various data sources and methodologies and to test these data in an integrated water resources modeling framework. Moreover, it is also the first time that large scale operational weather radar data has been implemented in hydrological applications in Denmark.

The main objective of this PhD study is to bridge the gap between radar meteorology and catchment hydrology by investigating the long term effects of using continuous radar based QPE data in a water resources model developed for the Skjern catchment. This primary goal is accomplished with special attention given to the following four aspects: the QPE algorithm issue, the uncertainty issue, the scaling issue, and the model calibration issue.

First of all, the radar QPE algorithm used in Denmark follows the guidelines reported by the BALTEX radar center (Michelson 2000). However, the developed radar QPE algorithm has never been documented and evaluated in details particularly from a combined meteorological and hydrological perspective. Furthermore, such an analysis requires that long time series are prepared.

Secondly, it has been extensively documented that using weather radar to estimate precipitation quantitatively could lead to sizeable uncertainties. To assess the impact of these uncertainties a practical method needs to be developed that quantifies the uncertainties in radar QPE while maintaining a reasonable computational time. An analysis on how the uncertainties in rainfall input data propagate through the hydrological modeling process is also needed.

Thirdly, simulation of hydrological fluxes driven by the change in precipitation is known to have strong scale dependency. It is crucial to examine the scaling issue under the influence of spatially distributed QPE data at various catchment sizes. Moreover, the scale dependency of the uncertainties in the hydrological simulations in response to the model input uncertainties also needs to be addressed.
Finally, model calibration plays an important role in redistributing water among different hydrological compartments. Therefore, the distributed water resources model considered in this study needs to be calibrated by using both radar and rain gauge based rainfall against field observations. The model performance with respect to the entire hydrological cycle should be compared between different models.

In the PhD thesis, a background introduction to radar QPE and its application in hydrology is given in Chapter 2. Chapter 3 outlines the scientific papers responding to the research objectives described above. The main conclusions and future prospective of the thesis subject can be found in Chapter 4. The four scientific papers, which summarize the main research activities during the PhD study, are included in appendices I - IV.
2. Background

The development of radar systems is accompanied by the development of radio technology, where the word ‘radar’ actually comes from the expression ‘RAdio Detection And Ranging’. Using radio as a way of communication, the radar technology went through evolutionary progresses during World War II. Soon after the war, many realized the potential of using radar for civilian purposes. As one of the most popular radar applications in our modern day life, radar is able to detect the location and track down the movement of weather systems. This chapter provides an overview on how the radar measures the weather targets and which methods that can be used to convert radar signals into precipitation (Battan 1973; Collier 1989; Rinehart 1997). The advantages and limitations of using radar technology in hydrological research will also be described.

2.1 Weather radar observation of the atmosphere

Radar is an active sensor that emits electromagnetic pulses into the surroundings. The backscattered energy that is reflected from the objects in its path is received by the radar. A typical radar system consists of at least the following four components: a transmitter that generates high frequency signals, an antenna that sends the signal out and receives the echoes returned, a receiver that processes the returned signals so that they are ready to be used, and a data display system (Rinehart 1997). The radar part that is visible to the general public is usually the antenna covered with a dome shield, a radome, and installed on top of an observation tower.

Weather radar is a ground based continuous remote sensing instrument. Both transmission and reception of the signals are carried out by the same antenna. In order to scan the entire atmosphere around the radar, the antenna first rotates horizontally and then moves to another pointing angle. After scanning of a number of angles, the whole volume scan of the atmosphere is completed.

The electromagnetic radiation in the nature has a very large range of frequencies, where only the band from 100 MHz to 100 GHz is normally used by the radar meteorologists. It was found convenient to designate letters to certain radar types based on frequency bands. Therefore the common weather radar systems can be classified as listed in Table 1. Pictures of commonly used radar types are shown in Figure 1.

Lower frequency and higher wavelength indicates that the radar has stronger signal power and less attenuation, therefore larger antenna dimension is required. As for weather observations, S-band WSR-88D radars are employed by the NOAA in the United States, whereas C-band radars are more commonly used by meteorological offices in Europe. The weather radar system discussed in the present PhD thesis is based on the EEC C-band radars operated by the Danish Meteorological Institute (DMI). Therefore, the theoretical explanations demonstrated below are also based on the characteristics of C-band radars. For other radar types, these are likely to vary.
Table 1. Radar bands and the corresponding frequencies and wavelengths (Rinehart 1997)

<table>
<thead>
<tr>
<th>Radar band</th>
<th>Frequency</th>
<th>Wavelength</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>1-2 GHz</td>
<td>30-15 cm</td>
</tr>
<tr>
<td>S</td>
<td>2-4 GHz</td>
<td>15-8 cm</td>
</tr>
<tr>
<td>C</td>
<td>4-8 GHz</td>
<td>8-4 cm</td>
</tr>
<tr>
<td>X</td>
<td>8-12 GHz</td>
<td>4-2.5 cm</td>
</tr>
<tr>
<td>Ku</td>
<td>12-18 GHz</td>
<td>2.5-1.7 cm</td>
</tr>
<tr>
<td>K</td>
<td>18-27 GHz</td>
<td>1.7-1.2 cm</td>
</tr>
<tr>
<td>Ka</td>
<td>27-40 GHz</td>
<td>1.2-0.75 cm</td>
</tr>
<tr>
<td>W</td>
<td>40-300 GHz</td>
<td>0.75-0.01 cm</td>
</tr>
</tbody>
</table>

Radar not only shows the location but more importantly the strength of a storm. This is because the calculated precipitation rate is related to the strength of the returned radar signal power. In order to use radar quantitatively, it is necessary to know the relation between the power received and the physical properties of the hydrometeor particles in the air (Villarini and Krajewski 2010). It is assumed that the rain drop particles are spherical and small compared to the radar wavelength. Thus,

\[ P_r = \frac{Z \cdot C \cdot |K|^2}{l^2 \cdot r^2} \cdot P_t \]  [1]

where \( P_t \) and \( P_r \) are the transmitted and received electromagnetic power, \( Z \) is the radar reflectivity factor (commonly called ‘reflectivity’), \( l^2 \) is the attenuation caused by the distributed media through which the radar signal propagates, \( |K|^2 \) is the dielectric constant, \( r \) is the distance from the target to the radar, and \( C \) is called the radar constant which can be expressed as below,

\[ C = \frac{\pi^2 \cdot g^2 \cdot \theta \cdot \phi \cdot h}{1024 \cdot \ln(2) \cdot \lambda^2} \]  [2]

where \( g \) is the antenna gain, \( \theta \) and \( \phi \) are the horizontal and vertical radar beam width, \( h \) is the pulse length and \( \lambda \) is the wavelength. All the terms on the right side of Eq.[2] are constant once a particular radar is built and therefore they are grouped as one constant factor. Knowing the radar power and some of the operational parameters, Eq.[1] can be easily rearranged and the radar reflectivity \( Z \) can be derived.
Figure 1. Commonly used weather radars (from left to right): a WSR-88D S-band radar operated by NOAA, the Danish Roemoe C-band radar, an X-band radar in Switzerland

2.2 Estimation of precipitation using weather radar

The radar reflectivity factor $Z$ is important to quantify because the reflectivity depends on the number and diameter of raindrops in the radar sample volume. Since the rain rate also depends on the number and diameter of the raindrops, the radar reflectivity and rain rate can be linked (Battan 1973). To demonstrate this theory mathematically, let’s take $N(D)$ as drop size spectrum with $D$ as the drop diameter. Then the reflectivity $Z$ (in mm$^6$/m$^3$) and rain rate $R$ (in mm/h) are given by,

$$Z = \int_0^\infty N(D) \cdot D^6 \ dD$$

and

$$R = \frac{\pi}{6} \int_0^\infty N(D) \cdot D^{3.67} \ dD$$

Eliminate $D$ by integrating Eq. [3] and [4], the following relationship can be obtained,

$$Z = A \cdot R^b$$

Eq. [5] is commonly called the Z-R relationship, which serves as the foundation for most of the operational radar QPE algorithms around the world. The most famous Z-R relationship is developed by Marshall and Palmer (1947) which is based on general stratiform precipitation.
Table 2. List of commonly used $A$ and $b$ values in the Marshall-Palmer equation (Battan 1973)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Drizzle</td>
<td>140</td>
<td>1.5</td>
</tr>
<tr>
<td>Widespread rain</td>
<td>250</td>
<td>1.5</td>
</tr>
<tr>
<td>Thunderstorm</td>
<td>500</td>
<td>1.5</td>
</tr>
<tr>
<td>Marshall-Palmer</td>
<td>200</td>
<td>1.6</td>
</tr>
<tr>
<td>DMI ARNE</td>
<td>220</td>
<td>1.6</td>
</tr>
<tr>
<td>WSR-88D</td>
<td>300</td>
<td>1.4</td>
</tr>
</tbody>
</table>

It should be mentioned that $A$ and $b$ are selected according to the type of precipitation (see Table 2). Thus, precise rainfall estimation requires the use of a dynamic Z-R relationship that accounts both for the differences between events and the changes as a precipitation system develops. However, in reality it is more common to use a fixed Z-R relationship for all situations.

In radar meteorology, a special unit is used for rainfall measurement. It is seen from above that the unit for radar reflectivity $Z$ is mm$^6$/m$^3$. In reality the $Z$ value has a tremendous range of six orders of magnitudes and therefore not convenient to use. Thus, it is much easier to use logarithmic scale instead of linear scale. The conversion can be seen below,

$$Z_{\text{log}} = 10 \cdot \log \left( \frac{Z_{\text{linear}}}{1 \text{ mm}^6 / \text{ m}^3} \right)$$  \[6\]

The new unit is called dBZ, which stands for decibels relative to reflectivity $Z$. Using the new unit, a 200 mm/h heavy storm corresponds to 60 dBZ based on Marshall-Palmer drop size distribution; whereas using the linear scale it is 1,000,000 mm$^6$/m$^3$.

2.3 Uncertainties in radar precipitation estimation

Rain gauges have been used for recording precipitation for centuries. Compared to the long history of the rain gauge, estimation of precipitation by radar is at its very early age. At the time when the radar QPE began to emerge, a discussion was undertaken on which one of the two measurements is more correct. For hydrologists, it is more common to consider the rain gauge measured rainfall as the ‘truth’ (Austin 1987; Joss and Lee 1995; Smyth and Illingworth 1998).

In fact, it is quite difficult to make a fair comparison since rain gauge and radar are measuring two different things. To begin with, radar measures precipitation at a certain height and rain gauge measures rainfall at the ground surface. Another big difference is the sampling area (Rinehart 1997). Taking the Danish situation as an example, the Hellman gauge has an opening area of 0.03 m$^2$, whereas the C-band radar has a sampling surface of 150 m long and 1000 m wide at approximately 60 km from the radar site, which corresponds to 150,000 m$^2$. Hence, the difference in measurement scale is 5 million times.
However, it is well known by the radar scientists that there are a large number of uncertainties associated with radar based rainfall estimation (Kitchen and Jackson 1993; Pamment and Conway 1998; Anagnostou et al. 1999; Vignal et al. 2000; Brandes et al. 2004; Holleman 2007). Some of them have been studied intensively while knowledge on others are still lacking. An illustration in Fig.2 summarizes most of the common radar uncertainty sources. In Denmark, the land surface is generally rather flat and no significant beam blockage is present, which can be seen as a big advantage for using radar technologies. However, other uncertainties still exist and will be briefly discussed below.

**Miscalibration**

In the present study the following terminology is used: ‘calibration’ is restricted to the radar hardware, whereas the word ‘adjustment’ is referred to the radar software. Miscalibration appears when the radar parts deteriorate with time and the radar constant ($C$ in Eq. [2]) is no longer valid. Using an incorrect $C$ value causes uniform bias of reflectivity within the radar footprint until it is recalibrated at a later time. Therefore, carefully calibration of the radar is crucial in establishing the Z-R relationship. DMI uses the NORDRAD method to calibrate the C-band radar which is described in DMI’s internal report (Gill and Overgaard 2005).

**Attenuation**

When the radar signal passes through a precipitation medium, the electromagnetic power is gradually reduced ($I^2$ in Eq. [1]). Attenuation is more serious for radars with shorter wavelength since the transmitted signal power is weaker, e.g., X-band radars have much larger attenuation problems than S-band radars. In addition, the density of the cloud also plays a role where denser cloud causes more signal attenuation. The attenuation correction method used at DMI is carried out by using the standard factor for water vapor, which is 0.17 dBZ/km.

![figure](image)

**Figure 2.** Sources of uncertainties that affect the accuracy of radar rainfall measurements (Rossa et al. 2005)
Anomalous propagation

When some special atmospheric conditions occur, such as sudden change in the air density, the radar beam is likely to bend downwards and hit the ground (see Fig. 2). Therefore the observed echoes are not real precipitation and are called radar clutters. It should be noticed that not only anomalous propagation can cause clutters; they also appear when sidelobes hit ground objects at short range or stationary obstacles presence close to the radar. For radar QPE, it is extremely important to remove spurious echoes, because they can greatly damage the data quality. Several methods have been proposed to deal with the clutter problems in Denmark, but they have not yet been implemented operationally (Boevith 2008).

Range degradation

Besides attenuation, other significant problems with radar estimation of precipitation at long range are beam broadening and overshooting (see also Fig. 2). Suppose a hypothetical situation where the radar beam has an opening angle of 1° and an elevation angle of also 1°, this will lead to a radar beam of around 2 km width at 100 km distance, thus causes partial beam filling. Moreover, at this distance the radar beams are almost 2 km above the ground, which is already higher than most of the precipitating clouds. To solve this problem, DMI uses a range dependent adjustment technique, short name ARNE, which is described in detail in (He et al. 2011).

Bright band

The technical name for this particular problem is called the variations in the vertical profile of reflectivity. It is known that the radar beams intercept with precipitation at different heights. Under some specific conditions, a layer of atmosphere is just below freezing point (usually around 3 km above sea level). When ice particles fall through this layer, they melt and eventually become rain. The melting process starts from the outside of the solid ice particles, and when the solid ice is coated with a film of liquid, it is extremely reflective, much more than either ice or water alone, seen on Fig. 3. Consequently it forms a ‘bright band’ on the radar display. The bright band contamination can be very harmful to radar QPE, but there hasn’t been a correction method implemented in Denmark.

Figure 3. Illustration of the transition from ice particles to raindrops (Pedersen 2009)
2.4 Merging of radar and rain gauge data

Radar has high spatial and temporal resolution in terms of areal precipitation estimation, whereas rain gauge offers better precision at point scale. As can be seen from above, using weather radar data alone embraces many types of uncertainties. Therefore, in most hydrological studies, spatial precipitation data from radar is normally combined with rain gauge data. The merging of radar and rain gauge data may be carried out using different degrees of complexity. The following lists some of the common methods.

Mean field bias correction

The assumption of this method is that the main radar uncertainty is caused by uniform bias across the radar field, or an incorrect parameter \( A \) in the Z-R relationship (Eq. [5]). Therefore the bias can be corrected by a multiplicative factor (Smith and Krajewski 1991),

\[
MFB = \frac{\sum_{i=1}^{n} G_i}{\sum_{i=1}^{n} R_i}
\]

where \( G_i \) is rain gauge observations over a certain area and \( R_i \) is the radar estimated values at the pixels that contain \( G_i \).

Range dependent adjustment

This method assumes that the main uncertainty from radar QPE is the distance degradation, such as attenuation, beam broadening and overshooting. A comprehensive range dependent adjustment method was developed during the Baltic Sea Experiment (Michelson and Koistinen 2000), and it was later adopted by DMI. It is assumed that the log scaled \( G/R \) is a function of the distance to the radar site \( (r) \), and can be expressed by a second order polynomial,

\[
\log(G / R) = a + br + cr^2
\]

where the dimensionless coefficients \( a \), \( b \) and \( c \) can be determined using least square curve fitting. A detailed description of the Danish implementation of the BALTEX method can be found in He et al. (2011).

Kriging with external drift

Instead of using radar as the primary data source, this method is based on the geostatistical interpolation of rain gauge data and uses radar data as the auxiliary information (Sinclair and Pegram 2005; Haberlandt 2007). It basically follows the same scheme as ordinary kriging except that an additional constraint is added using the rainfall estimation from radar,

\[
\sum_{i=1}^{n} \lambda_i \cdot R_i = R_0
\]
where $R_i$ is the radar pixel value at the rain gauge location $i$, $\lambda_i$ is the weight for $R_i$, and $R_0$ is the radar value at the estimation location. Eq. [9] is solved together with the coupled linear equations during the kriging process. It is seen that this method is much more computationally demanding than the other methods described above.

**Probability matching**

This method is different from the rest of the radar - rain gauge merging techniques in that it does not use a Marshall-Palmer like Z-R relationship. Therefore, instead of using the physical properties of the rain drop size distribution, the probability matching model is purely based on statistics (Rosenfeld et al. 1994). It is assumed that reflectivity $Z$ measured by radar and rain rate $R$ measured on the ground are related as below,

$$\int_0^R P(R)dR = \int_0^Z P(Z)dZ \tag{10}$$

In this case $Z$ and $R$ are both random variables and $P(x)$ is an unconditional probability distribution function. Therefore, the probability distributions of $Z$ and $R$ over a time and space match the empirical method.

### 2.5 Application of radar meteorology in hydrology

The awareness of using weather radar in hydrology didn’t emerge until after 1990s. With the advances of the radar technology, a growing number of hydrologists realized that using radar in hydrological applications has a promising future (Sumner 1990; Ramsey 1995; Peters and Easton 1996; Andrieu et al. 1997; Carpenter et al. 1999).

Among the many hydrological applications employing radar, the most significant contribution came from the estimation of extreme rainfall (Vieux and Bedient 1998; Collier 2007; Habib et al. 2009), sometimes defined as precipitation over 500 mm in 24 hours or over 100 mm in 1 hour. Extreme rainfall conditions often occur at local scale and it was found that radar can provide valuable information of the rainfall fields where the scale of the storm is less than the spacing between rain gauges.

Following this idea, the high resolution radar data gathered in real time has become increasingly important for operational flood forecasting. Several case studies have shown improvements in the description of precipitation fields measured by radar which allows more advanced designs of the flood warning system. Those designs were not practical previously where the information on rainfall was entirely based on rain gauge (Krajewski and Smith 2002; Morin et al. 2006; Borga et al. 2007; Gourley et al. 2010).
Using weather radar can also be beneficial to the design of urban drainage systems. Radar is able to provide near future quantitative precipitation forecast (QPF) for the next few hours, also known as ‘nowcasting’. Therefore, by using the knowledge from the nowcast, valuable time is earned for adjusting the wastewater treatment plant, since it takes time to switch from dry to wet mode. QPF can also facilitate the operation of the sewage water routing, where the water is dispatched from areas with overload to areas that are less affected by the storm (Smith and Krajewski 1991; Bell and Moore 1998; Tilford et al. 2002).

The large area coverage and high spatial resolution also makes radar based precipitation products very suitable for distributed hydrological modeling on water resources. In principle, radar is able to provide rainfall measurements down to the model grid scale whereas the rain gauge based products are simply interpolations to each model grid using a limited number of observation points. Using radar QPE data for such models is the main focus of the present PhD research.
3. Introduction to thesis papers

The PhD thesis consists of four research articles (Paper I-IV). An outline is given to facilitate the understanding of the interconnections between the articles. At the initial stage of the PhD study, one year of radar reflectivity data and a package of QPE software codes, ARNE, was provided by DMI. Therefore, the focus of Paper I was to describe the methodology of the ARNE algorithm and make the first effort to use such data in hydrological modeling. Since the available data wasn’t sufficient to carry out a dedicated model calibration, the hydrological model applied in Paper I was calibrated using rain gauge based rainfall data. While Paper I was finished, the radar data uncertainty issue began to emerge. As a result, the work on Paper II was initiated where a multigaussian simulation approach was tested for quantification of the radar QPE uncertainties. Subsequently, the hydrological consequences caused by using the rainfall ensemble were also analyzed. At the last period of the PhD, a long term data series was ready to be tested in a hydrological modeling framework. Finally, Paper III dealt with the model calibration issues using radar data as precipitation input. Moreover, as a sideline of the radar precipitation, contribution was also made to produce gridded rainfall maps by interpolating rain gauge data in Paper IV.

Across the four papers, consistencies can be found in a number of tools and methodologies used. All research activities are carried out in the Skjern catchment which is located in western Denmark. The study catchment, which covers an area of approximately 3500 km$^2$, is equipped with state-of-the-art hydrological observation instruments (Jensen and Illangasekare 2011). All the radar QPE data used in the studies were initially countrywide rainfall estimations and then selected for the area of interest. All the hydrological modeling practices were performed based on the MIKE SHE code, which in the present study is used as a deterministic, distributed and physically-based modeling system (Abbott et al. 1986).

3.1 Introduction to Paper I

Title: An operational weather radar based QPE and its application in catchment water resources modeling

The Danish Meteorological Institute operates a radar network consisting of five C-band Doppler radars. Radar based QPE is performed on a daily basis and is considered to have the potential to significantly improve the spatial representation of precipitation compared to rain gauge based methods. Thus, it may provide the basis for better water resources assessments. In Paper I a radar QPE algorithm termed ARNE is presented. ARNE is a distance dependent areal estimation method that merges radar data with ground surface observations. The method is applied to the Skjern catchment for the year of 2006, where alternative precipitation estimates are used as input to the hydrological model. The hydrological responses from the model are analyzed by comparing radar and ground based precipitation input scenarios. The results show that radar QPE products are able to generate reliable simulations of stream flow and water balance. The potential of using radar based precipitation is found to be especially high at smaller scale where the impact
of spatial resolution is evident from the stream discharge results. Also groundwater recharge is shown to be sensitive to the rainfall product selected.

3.2 Introduction to Paper II

Title: Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling

Uncertainty analysis in hydrological modeling has become an essential step in the scientific interpretation of model results. Among many uncertainty sources in the modeling practice, precipitation estimation plays an important role since it is the main driving force for other hydrological processes. Paper II demonstrates a statistical method for generating radar rainfall realizations that account for the uncertainties in radar QPE. The random sampling technique used to generate stochastic uncertainty fields is based on Sequential Gaussian Simulation. The hydrological impact of the uncertainties in radar QPE is analyzed by propagating the rainfall ensemble through a distributed water resources model. The study shows that the uncertainty of the simulated stream discharge is closely related to the intensity of the rainfall input signal. The coefficient of variation is calculated for simulated stream discharge and groundwater recharge at subcatchments with various sizes. The results reveal strong scale dependency showing higher variations of hydrological uncertainties at smaller catchments. The uncertainties from precipitation input are generally amplified in the hydrological model. The effect is less significant for groundwater recharge but rather substantial for stream discharge.

3.3 Introduction to Paper III

Title: Calibration and validation of a distributed water resources model using radar and rain gauge based precipitation input

Radar QPE has high spatial representativeness of the rainfall systems and is therefore very suitable in distributed hydrological modeling of water resources. Paper III investigates the potential of using long term radar QPE data for a distributed water resources model. The work provides valuable insights that are difficult to obtain if only event based data are used. The study is carried out using a 5-year radar QPE data series to calibrate and evaluate a hydrological model developed for the Skjern catchment. Comparisons are made with simulation results by a model calibrated using only rain gauge data. The model calibration suggests that the root depth of the crops is the most sensitive parameter to the different rainfall products, since it determines the evaporated water fluxes in order to achieve the water balance. Evaluation of the optimized models shows that the radar based models generate comparable results to the rain gauge based model in terms of stream discharge, groundwater head and overall water balance. However, if realistic rainfall estimates should be produced it is not sufficient to correct the radar data only for the areal mean bias. Hence, an indicated range dependent adjustment method is preferred. The results also suggest that the advantages of high resolution data can be hampered by the uncertainties in the QPE, especially the problem caused by the accumulated errors of radar QPE. Furthermore, an attempt is made to validate the spatial patterns generated by the hydrological model using satellite based surface temperature data.
3.4 Introduction to Paper IV

*Title: Impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resources model*

Precipitation is a key input variable to hydrological models. Its spatial variability is expected to impact the hydrological response predicted by a distributed model. In Paper IV, the effect of spatial resolution of precipitation on runoff, recharge and groundwater head is analyzed in the Alergaarde catchment, which is an upstream subcatchment within the Skjern catchment. Six different precipitation spatial resolutions are used as inputs to a physically based hydrological model. The results show that the resolution of precipitation input has no apparent effect on the annual water balance and the runoff hydrograph at the watershed outlet. However, groundwater recharge and groundwater heads are both affected. The impact of the spatial resolution of precipitation input becomes more noticeable for decreasing catchment size. The highest resolution of 500 m was found to result in the best representation of the hydrological response at subcatchment scale. Furthermore, stream discharge, groundwater recharge, and groundwater head are affected by the method that systematic errors in precipitation measurements are corrected. Therefore, it is suggested that as long as the average precipitation of the considered catchment or subcatchment is accurately estimated, the spatial resolution seems less important when the integrated response in the form of stream flow is considered.
4. Conclusions and perspectives

A methodology for generating daily radar QPE product by using the ARNE algorithm has been introduced. Evaluations carried out for 2006 suggest that using raw radar data without adjustment could cause severe underestimation of rainfall. However, ARNE was able to successfully remove both systematic mean bias and distance induced bias. It is also shown that that overall quality obtained by merging radar and rain gauge measurements can be superior to the case where single data source is used.

The main contribution of the PhD study is the application of radar QPE for long term water resources modeling. The outputs from the calibrated hydrological models using both radar and rain gauge based rainfall data have been analyzed. The results suggest that using QPE data as rainfall input, hydrological models are able to generate reliable simulations of surface water, groundwater, and total water balance. The simulated stream discharge is found favorable in 2006 when the quality of the QPE data is high. Simulated groundwater head shows very similar results among the models tested, with small preference given to the radar based models. The simulation of total water balance suggests that all models perform equally well despite some differences in the internal flow paths.

In the PhD study, a statistical method has been developed for generating radar rainfall ensembles as input to the Skjern model. Two case studies covering different rainfall regimes demonstrate that a minimum number of 200 realizations are required to properly represent the error structure and also keep the computational time for hydrological simulations acceptable. When uncertainties of long term radar QPE data are propagated through the hydrological model, the confidence band has similar dynamics compared to the observed discharge but fails to include more than half of the observation points. In addition, it is discovered that the impact of radar rainfall uncertainties is inversely correlated to the catchment size.

The hydrological model shows that relatively small subcatchments with an area less than 400 km$^2$ are highly sensitive to the spatial rainfall distribution, while for larger areas it is more dependent on the aggregated precipitation volume. The scaling issues are further investigated by dividing the catchment into 148 subcatchments with sizes smaller than 50 km$^2$. The simulated groundwater recharge has the same degree of variation compared to the rainfall input signal, but for simulated stream discharge the variation is magnified by a factor of three.

The study indicates that the quality of the radar based precipitation play an essential role for the performance of the hydrological model. Hence, in order to improve the use of QPE for hydrological purposes, the uncertainty on the raw radar data should be kept as low as possible, which indicates that programs for quality assurance, such as removal of residual bias and the random clutters, should be implemented. Therefore in an outlook to the future prospects, the most urgent task is to improve data quality in radar measurement.
It can be also very interesting to operate radar QPE on a sub-daily basis, since the advantage of temporal precision offered by radar hasn’t been investigated by the present study. Moreover, considering that the study area is nearly outside the optimal detection range of the C-band radars, an X-band radar installed locally within the Skjern catchment can be very beneficial for hydrological purposes. Finally, new radar technologies, such as using antenna with dual-polarization capacity, may also assist to increase accuracies in the rainfall estimation.


AN OPERATIONAL WEATHER RADAR BASED QPE AND ITS APPLICATION IN CATCHMENT WATER RESOURCES MODELING

Xin He, Flemming Vejen, Simon Stisen, Torben O. Sonnenborg, Karsten H. Jensen

Published in Vadose Zone Journal, January 2011
An Operational Weather Radar–Based Quantitative Precipitation Estimation and its Application in Catchment Water Resources Modeling

The Danish Meteorological Institute operates a radar network consisting of five C-band Doppler radars. Quantitative precipitation estimation (QPE) using radar data is performed on a daily basis. Radar QPE is considered to have the potential to significantly improve the spatial representation of precipitation compared with rain-gauge-based methods, thus providing the basis for better water resources assessments. The radar QPE algorithm called ARNE is a distance-dependent areal estimation method that merges radar data with ground surface observations. The method was applied to the Skjern River catchment in western Denmark where alternative precipitation estimates were also used as input to an integrated hydrologic model. The hydrologic responses from the model were analyzed by comparing radar- and ground-based precipitation input scenarios. Results showed that radar QPE products are able to generate reliable simulations of stream flow and water balance. The potential of using radar-based precipitation was found to be especially high at a smaller scale, where the impact of spatial resolution was evident from the stream discharge results. Also, groundwater recharge was shown to be sensitive to the rainfall product selected. Radar QPE appears to have unprecedented potential in optimizing precipitation input to distributed hydrologic models and thus model predictions.

Abbreviations: CAPPI, Constant Altitude Plan Position Indicator; DMI, Danish Meteorological Institute; MFB, mean field bias; QPE, quantitative precipitation estimation.

Weather radar is a sophisticated remote sensing instrument that measures the reflectivity of objects in a given volume of the atmosphere. Research in the use of weather radar based quantitative precipitation estimation (radar QPE) in hydrologic applications has increased in recent years mainly due to the increasing demand for more refined spatial and temporal resolution of rainfall products for modeling purposes. Many meteorological institutes around the world, including the Danish Meteorological Institute (DMI), have considered using radar QPE as an important supplement to the conventional rain gauge rainfall products (Klazura and Imy, 1993; Fulton et al., 1998; Golding, 2000; Harrison et al., 2000; Tabary, 2007; Tabary et al., 2007).

Operational radar QPE is a complex system that involves several elements including hardware design, signal processing, image analysis, data quality control, uncertainty analysis, and database organization. Due to the fact that each radar operation has its unique climatological conditions, geographical terrain, and economic capacity, however, agreement on a standard procedure in the development of radar QPE products can hardly be reached.

Despite the lack of standard procedures in implementation, the basic principles of radar precipitation estimation are well known and have been discussed in many textbooks (Batton, 1973; Collier, 1989; Rinehart, 1997). Radar emits and measures electromagnetic waves backscattered by raindrops; these electromagnetic waves are directly related to reflectivity. The relation between radar reflectivity (Z) and precipitation rate (R) is physically based because both Z and R are integrations of the raindrop size distribution (DSD); however, the Z–R relationship can be described by many different approaches. One of the common approaches is to deploy distrometers to observe the DSD at ground level. The Z–R relationship can thereby be computed by statistical analysis of the DSD observations and selected according to different rainfall regimes (Steiner and Smith, 2000; Ulbrich and Miller, 2001; Brinig et al., 2004; Lee and Zawadzki, 2006). Another approach is to establish the Z–R relationship between radar reflectivity and surface rainfall measurement.

A methodology for quantitative precipitation estimation (QPE) was developed. The cumulative effects of radar QPE on hydrologic predictions at a long time scale were assessed. The results show that the influence of radar QPE on the accuracy of a designated hydrologic model is highly sensitive to the size of the catchment.

X. He and K.H. Jensen, Dep. of Geography and Geology, Univ. of Copenhagen, Oster Voldgade 10, DK-1350, Copenhagen, Denmark; X. He and F. Vejen, Data and Climate Division, Danish Meteorological Institute, Lyngbyvej 100, Copenhagen, Denmark; and S. Stisen and T.O. Sonnenborg, Dep. of Hydrology, Geological Survey of Denmark and Greenland, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark. *Corresponding author (xh@geo.ku.dk).
The main reason that radar QPE may not represent surface rainfall properly is that most radar beams are far above the ground. Differences between ground rainfall estimates derived from radar and rain gauges can be categorized from different perspectives. If the error characteristics are of interest, they are commonly classified as (i) systematic bias that is a result of using a uniform \( Z-R \) relationship to represent a varying \( Z-R \) relationship, attenuation by a wet radome, or radar calibration problems; (ii) range-dependent error that is mainly due to the vertical variation in the reflectivity profile with distance, attenuation of the radar pulse by rain, or radar beam occultation; and (iii) random errors from radar signal noise such as radar clutter effects. If the classification is performed from an operational point of view, errors in the measurement of reflectivity or in the process of retrieving the rainfall estimation using the \( Z-R \) relationship may appear (Collier et al., 1983; Kitchen and Jackson, 1993; Klazura et al., 1999; Germann et al., 2006; Ciach et al., 2007).

Due to the lack of agreement between ground rainfall estimates based on elevated radar reflectivity and those based on rain gauges, algorithms must be developed to convert radar signals to rainfall intensity at ground level. Rain gauges provide point measurements with a precision that is still considered to be the most accurate at ground level. Therefore, the common solution to obtain an optimal rainfall product is to combine radar and rain gauge information such that the strengths from both products are utilized. Methods for merging radar and rain gauge data have many degrees of complexity, starting with a simple bulk adjustment in mean field bias (Smith and Krajewski, 1991). More complex methods have been proposed such as co-kriging (Krajewski, 1987; Sun et al., 2000), kriging with external drift (Sinclair and Pegram, 2005; Haberlandt, 2007), or Kalman filtering (Seo and Breidenbach, 2002; Chumchean et al., 2006). To implement those complex methods operationally is very time consuming and computationally demanding; therefore, the choice of a radar–rain gauge merging technique is subjective and largely depends on the radar operator’s experience and practical conditions (Goudenaufdt and Delobbe, 2009; Velasco-Forero et al., 2009). The current radar QPE system in use in Denmark was used in this study and suggestions were developed to improve the quality of the radar QPE products; however, a thorough description of the error correction strategies is beyond the scope of this work.

It is commonly acknowledged that precipitation plays an important role in the simulation of the water balance and hydrologic behavior at the catchment scale (Beven, 2001, 2002). The sensitivity of hydrologic models to precipitation resolution in space and time has been studied in many experiments (Watts and Calver, 1991; Finnerty et al., 1997; Singh, 1997; Winchell et al., 1998; Koren et al., 1999; Bell and Moore, 2000; Carpenter and Georgakakos, 2006; Morin et al., 2006). All these studies have generally concluded that the spatial and temporal resolution of rainfall input to a model can significantly influence the timing and peak of the simulated catchment response. Although rain gauge rainfall products are still considered to be a reliable data source, increasing attention has been given in recent years to the use of radar precipitation in forward hydrologic modeling, where comparisons are made based on model performance using radar and rain gauge data (Bell and Moore, 1998a,b; Sun et al., 2000; Wood et al., 2000; Carpenter et al., 2001; Carpenter and Georgakakos, 2004; Guo et al., 2004; Cole and Moore, 2008, 2009). Inverse hydrologic modeling involving radar QPE has also been attempted in several studies (Corral et al., 2000; Borga, 2002; Morin et al., 2009). Most studies on the hydrologic applications of radar QPE have been concerned with flood forecasting and surface water management; however, impacts on other hydrologic processes and the overall water balance have not been fully investigated. Moreover, the application of radar data is normally constrained to a short time scale, of hours or days. The pros and cons regarding long-term radar QPE in water balance modeling, especially the difference between accumulated radar precipitation and traditional methods of rainfall measurement, have not been intensively studied.

The main objectives of this study were (i) to document and investigate the Danish radar QPE algorithm from both meteorological and hydrologic perspectives, and (ii) to analyze the impact of different precipitation input scenarios on the simulation of water fluxes and the water balance of the Skjern catchment using a distributed integrated hydrologic model.

**Materials and Methods**

**The Danish Radar Quantitative Precipitation Estimation Algorithm**

The first radar deployed in Denmark for weather surveillance purposes was installed in 1986 to serve the Copenhagen airport. Since then the radar network has been constantly updated. By 2009, the DMI has been responsible for operating five C-band Doppler radars that offer a countrywide coverage. The geographic locations of the radars are shown in Fig. 1. All of the radars were purchased from Electronic Enterprise Corporation within the past few years. The most recent installations at Bornholm (in 2007) and Virring (in 2008) are capable of making dual-polarization scans. The operational radar parameter settings are summarized in Table 1. Volume scans are undertaken every 10 min with a maximum detection range of 240 km. With the software provided by the
radar manufacturer, the Constant Altitude Plan Position Indicator (CAPPI) products are visualized at heights from 1 to 12 km at 1-km vertical intervals. It is a common technique for radar operators to produce an artificial CAPPI image at a certain height by using higher elevation angles to fill the near ranges and the lower elevation angle to fill the far ranges so there will be no blank areas on the integrated CAPPI image (so-called pseudo-CAPPI). The DMI uses software developed in-house to generate pseudo-CAPPIs at 2 km above ground level with the same time frequency and a horizontal resolution of 2 km.

The Danish radar QPE algorithm (called ARNE) is based on the method developed by the Baltic Sea Experiment (BALTEX), which is a European regional project focusing on the hydrologic cycle and the exchange of energy between the atmosphere and the Earth’s surface (Michelson, 2000; Michelson and Koistinen, 2000). A flow chart illustrating the overall processing chain of ARNE is shown in Fig. 2.

Taking the main concept from the BALTEX method, ARNE has been modified to accommodate the Danish climatological and instrumental conditions from a number of aspects. First, retrieval of pixels from the raw radar images has several resampling options, which include averaging among nine pixels in the vicinity and bilinear resampling. Second, considering the relatively flat terrain in Denmark, the composite strategy of radar images is either to choose the nearest radar antenna or the maximum output values in the overlapping area. Other countries in the BALTEX region are characterized by mountainous terrain and the composite scenario was therefore originally developed to obtain values from the lowest radar beam. Third, a 7-d moving window is used to obtain radar \((R)\) and rain gauge \((G)\) observation pairs, which means ARNE is performed by considering \(G–R\) samples outside the current day. It is specified that the minimum number of \(G–R\) pairs has to be \(>200\) in the moving window. This empirical number is set to avoid the risk that the number of samples during the past 7 d with precipitation may not be sufficient to perform the distance corrections based on the climatological situation in Denmark. Last but not least, the accumulation period is 24 h instead of the 3 and 12 h used in the BALTEX method, so that errors caused by the vertical profile of reflectivity will be largely smoothed out.

According to the description above, ARNE is regarded as a range-dependent accumulated precipitation estimation algorithm. Its main scripts have three primary components: (i) calculation of the accumulated precipitation from the radar, (ii) establishment of relationships between radar measurement bias (obtained from comparing with rain gauge observations) and distance (from the radar site to the rain gauge sites), and (iii) generation of spatial adjustment factor fields. These components are explained in detail below. To generalize the estimation of radar QPE, assumptions are made that the rain gauge measurements are ground truth, that the radar measurements account for the true spatial and temporal variability of the areal precipitation field, and that the radar and rain gauge measurements are both valid at the same location in time and space.

Table 1. Operational parameter settings for the Danish national radar network.

<table>
<thead>
<tr>
<th>Radar parameters</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>5 cm (C-band)</td>
</tr>
<tr>
<td>Range</td>
<td>240 km</td>
</tr>
<tr>
<td>Pulse length</td>
<td>2 (\mu)s</td>
</tr>
<tr>
<td>Pulse repetition frequency</td>
<td>250 Hz</td>
</tr>
<tr>
<td>Peak power</td>
<td>250 kW</td>
</tr>
<tr>
<td>Rotation rate</td>
<td>20 °/s</td>
</tr>
<tr>
<td>Antenna gain</td>
<td>45 dB</td>
</tr>
<tr>
<td>Vertical beam width</td>
<td>±0.5°</td>
</tr>
<tr>
<td>Horizontal beam width</td>
<td>1.0°</td>
</tr>
<tr>
<td>Volume scan angles</td>
<td>0.5, 0.7, 1.0, 1.5, 2.4, 4.5, 8.5, 13.0, 15.0</td>
</tr>
<tr>
<td>Scans</td>
<td>plan position indicator</td>
</tr>
</tbody>
</table>
According to Michelson (2000), the accumulated precipitation is first estimated based on the 10-min pseudo-CAPPI images kept in the DMI’s radar archive as radar reflectivity (in dBZ). These images are piped into ARNE and converted to rainfall intensity (mm/h) using the classic Marshall–Palmer equation (Marshall et al., 1947; Marshall and Palmer, 1948):

$$Z = AR^b \rightarrow R = \left(\frac{Z}{A}\right)^{1/b}$$  \[1\]

with standard parameter values $A = 220$ and $b = 1.6$ ($A$ and $b$ are both dimensionless). Estimated 10-min rainfall intensities are accumulated as 24-h sums. Daily rain gauge point observations are then matched with pixels on the radar rainfall sums that have the same geographic coordinates.

Second, the relationship between the distance from the radar and the bias is established. The bias of radar measurements is represented by the logarithmic gauge/radar ratio $F_g = \log(G/R)$ for each collocated radar–gauge pair. Based on the available $G–R$ pairs, the distance dependency of the bias can be formulated for every pixel on the radar composite. Based on the assumption that a second-order polynomial will enable a nonlinear relationship to reflect the systematic variation of reflectivity with distance, the fundamental methodology is to use a second-order polynomial to express the relation between bias and distance:

$$F_{r(i,j)} = a + br + cr^2$$ \[2\]

where $F_{r(i,j)}$ is the distance-dependent adjustment factor field and $r$ is the distance to the radar, while coefficients $a$, $b$, and $c$ (dimensionless) can be determined using least square fits between all $F_g$ values and their corresponding distances to the radars. The $F_{r(i,j)}–r$ relationship provides the estimate of the bias ($\hat{F}_g$) at each rain gauge location. A quality control is then conducted to evaluate the difference between the observed and estimated bias, $F_g$ and $\hat{F}_g$, respectively, to determine whether the observed bias is captured by the model:

$$Z_g = \frac{F_g - \hat{F}_g}{\sigma}$$ \[3\]

where $Z_g$ is the normalized residual and $\sigma$ is the standard deviation of $F_g$. A quality flag, $Q_{g}$, is assigned to each value of $Z_g$ using the following criteria:

$$Q_g = \begin{cases} 1 & \left|Z_g\right| \leq \sigma \\ 1 - \left|Z_g\right| & \sigma < \left|Z_g\right| \leq 2\sigma \\ 0 & \left|Z_g\right| > 2\sigma \end{cases}$$  \[4\]

Observations with $Q_g = 0$ are rejected and a new $F_{r(i,j)}–r$ relationship can be established using the quality-ensured $G–R$ pairs. The
The spatial weighting factor of Eq. [5], $w_{s(i,j)}$, is determined by

$$w_{s(i,j)} = \exp \left(-\frac{D_{p(i,j)}}{D_0}\right)$$  \[10\]

where $D_0$ is the decorrelation distance, which is based on an analysis of the spatial correlation structure of the $F_g$ data set. The value of $D_0$ characterizes the distance between gauges where the correlation between errors is insignificant. It is assumed that the decorrelation distance for each day is constant across the entire rainfall field because all qualified rain gauges in the country are taken into account. The value of $D_0$ is obtained by using $F_g$ values for each gauge as well as their distance from each other. For each 10-km interval, the semi-variance ($\gamma^2$) between observations is calculated. The value of $D_0$ is estimated as twice the distance where $\gamma^2 = 1/e$, assuming a linear interpolation between two quantized distance intervals. The variable $D_p(i,j)$ is the local observation density, defined by the square root of the area covered by five local $F_g$ observations. The observation density is not constant so that the weighting factor has to vary accordingly to account for the local variations (Koistinen and Puhakka, 1981).

With all these steps finished, the adjustment of the original radar composite sum, $R_{r(i,j)}$, becomes

$$R_{r(i,j)} = R_{i,j}10^{F_g(i,j)}$$  \[11\]

where $R_{r(i,j)}$ is the final adjusted daily rainfall field.

As an alternative to this radar–rain gauge merging technique, a simple mean field bias (MFB) correction method was used in this study for comparison purposes. The MFB is defined as

$$\text{MFB} = \frac{\sum_{i=1}^{n} G_i}{\sum_{i=1}^{n} R_{bi}}$$  \[12\]

and

$$R_{i,j} = R_{b(i,j)} \times \text{MFB}$$  \[13\]

where $G_i$ is the $i$th rain gauge observation and $R_{b(i,j)}$ is the radar-estimated value at the pixels that contain $G_i$. With $R_{b(i,j)}$ being the original radar precipitation field, the adjusted precipitation field [$R_{i,j}$] suggests that the entire precipitation field will be multiplied by a single correction factor regardless of the distance relations. Therefore the original radar rainfall spatial structure will be preserved, whereas ARNE reshapes the rainfall field based on rain gauge locations. It should be mentioned that unlike the method in ARNE that uses a 7-d moving window, MFB takes into account only the radar and rain gauge data from the current day.
Grid Precipitation Products from the Danish National Rain Gauge Network

The present Danish national rain gauge network consists of three types of rain gauges: the manual Hellmann gauge, the Geonor automatic gauge, which measures precipitation based on the vibrating string principle (Bakkehoi et al., 1985), and the Rimco 7499 tipping bucket gauge. The Hellmann gauges have been deployed nearly 100 yr and were the most widely used rain gauges in Denmark in 2006. The first automatic rain gauge was introduced to Denmark in 1979. With the update of rain gauge networks over the years, automatic gauges are concentrated around the large cities, leaving the countryside dominated by manual gauges. The temporal resolution of rainfall measurement for Hellmann, Geonor, and Rimco gauges are 24 h, 10 min, and 1 min, respectively. The Danish national rain gauge network in 2007 is presented in Fig. 3a.

It is known that wind-induced errors in rain gauge precipitation measurement will result in systematic deficits, especially at high wind speed. The magnitude of the deficit can be as high as half of the true precipitation. To account for this error, a systematic error correction method was developed, where a multiplicative factor is applied to each observed rain gauge recording (Allerup et al., 1997). Because the correction factors consider the architectural and vegetation surroundings close to the gauge, rain gauge catch correction is performed by using wind shelter categories (A, B, and C) that are assigned to each rain gauge with seasonal variations. The correction factors are provided on a monthly basis and are shown in Table 2. All rain gauges are corrected based on this table before the measurements are subjected to any type of application. Uncertainty in the calculation of correction factors has been found to be approximately 4% in terms of standard deviations. The factors are based on data from the period 1961 to 1990. The average precipitation amount in the recent period has, however, been significantly higher than in the calculation period. Therefore the correction factors listed here are prone to a larger degree of uncertainty if applied to recent situations.

The grid precipitation products based on rain gauge information were represented using two techniques in this study. One is DMI’s 10- by 10-km grid product (DMI10 in short), which uses an inverse distance technique (Scharling et al., 2006). The other technique uses ordinary kriging, which is one of the most widely used interpolation methods in geoscience to estimate point values at locations without observation by weighting of neighboring points (Davis, 1986).

Rain gauge precipitation products are included in this study for various reasons. The results and application of the radar QPE should be compared and validated against the most reliable data sources. The DMI10 is recognized as one of the official Danish rainfall products, which has been used for developing and

---

**Table 2. Monthly rain gauge catch correction factors for Denmark recommended by the Danish Meteorological Institute (Allerup et al. 1997).**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Freely exposed (C)</td>
<td>1.53</td>
<td>1.53</td>
<td>1.45</td>
<td>1.29</td>
<td>1.16</td>
<td>1.13</td>
<td>1.12</td>
<td>1.13</td>
<td>1.17</td>
<td>1.29</td>
<td>1.48</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td>Moderate wind (B)</td>
<td>1.41</td>
<td>1.42</td>
<td>1.35</td>
<td>1.24</td>
<td>1.13</td>
<td>1.11</td>
<td>1.10</td>
<td>1.11</td>
<td>1.14</td>
<td>1.23</td>
<td>1.37</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>Ideal wind (A)</td>
<td>1.29</td>
<td>1.30</td>
<td>1.26</td>
<td>1.19</td>
<td>1.11</td>
<td>1.09</td>
<td>1.08</td>
<td>1.09</td>
<td>1.10</td>
<td>1.17</td>
<td>1.26</td>
<td>1.16</td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 3.** The Danish national rain gauge network, including three types of rain gauges: (a) operational rain gauges up to present time, used for radar-gauge adjustment, and (b) rain gauges that had observations until 2006 and were closed down in 2007. This group was subjected to validation.
calibrating distributed hydrologic models in Denmark in the past. The DMI10, however, is based on all available rain gauge records. The intention of the present study was that part of the rain gauges should be reserved for validation purposes. Additionally, the grid size of the DMI10 is not consistent with the radar QPE products. Therefore, kriging was introduced not only to provide an alternative rain gauge rainfall product but also because it is easy to control the number of rain gauges used in spatial interpolation and grid size on the rainfall domain. The kriging domain had a spatial resolution of 2 by 2 km and it was thus consistent with that of the radar rainfall domain and the calculations were performed by using only data from the rain gauges shown in Fig. 3a.

Hydrologic Model for the Skjern Catchment

The Skjern catchment is located on the west coast of the Jutland peninsula. The catchment covers an area around 3500 km² excluding the fjord area, with land use consisting of 56% agriculture, 29% grass, 5% heath, 7% forest, and 2% urban areas. The shallow geology in western Jutland is dominated by Quaternary outwash plains of sand and gravel. Isolated islands of Saalian sandy till are found between the outwash plains. The thickness of the Quaternary deposits varies between 50 and 250 m. Miocene sediments formed by alternating layers of clayey and sandy marine deposits are found below. These sediments have a thickness of 200 to 300 m. The Quaternary and Miocene sand formations form large interconnected aquifers (van Roosmalen et al., 2007). Precipitation is measured by more than 20 rain gauges located across the catchment (Fig. 4). Records show that the annual average precipitation is 1057 mm, with a 50% variation during the years 1990 to 2004 (van Roosmalen et al., 2007). Stream discharge gauging stations are also shown in Fig. 4.

A water resources model was developed for the Skjern catchment with the intention to explore the entire water cycle and the water balance at the catchment scale. The MIKE-SHE code originally derived from the SHE model (Abbott et al., 1986) was used as the modeling framework. The MIKE-SHE code is a deterministic, distributed, physically based hydrologic modeling system that integrates the entire land phase of the hydrologic cycle, including surface and groundwater. The model for the Skjern catchment was built on a 500- by 500-m grid and was based on the current version of the Danish National Water Resources Model, the DK-Model (Henriksen et al., 2003). New field data collected as part of the HOBE project will be integrated into the model and it is an ideal tool for investigating water balance issues while experimenting with new modeling strategies (Jensen and Illangasekare, 2011; Stisen et al., 2011).

We investigated the impact of different precipitation input scenarios on water balance modeling with particular emphasis on the radar QPE products. The baseline Skjern River model was calibrated by the PEST optimization tool (Doherty, 2004) using observed stream discharge and hydraulic head data. Rain gauge data were used to define the precipitation input, and Thiessen polygons were used to estimate the spatial rainfall distribution. In the calibration process, the number of free parameters was limited to nine, which is relatively few considering the substantial number of parameters in the model setup. More details on the objective functions as well as their weights can be found in Stisen et al. (2011). For the hydrologic simulations using various kinds of precipitation inputs, the model was run for 2002 to 2006. Only the precipitation for the last year was replaced each time by different data sources. Hence 2002 to 2005 can be regarded as the warming up period to ensure that the different precipitation scenarios were tested for the same initial condition in the beginning of 2006.

Results and Discussion

Precipitation Estimation Based on Radar and Gauge Data

A case study was undertaken based on the precipitation estimation schemes described above. Both radar and rain gauge data are presented in daily values. The date of 27 June 2006 was chosen as an example because it represented a typical rainfall event in Denmark based on the structure and intensity of the rainfall. The meteorological condition of this case was frequent showers and occasional heavy convective cells. A total of 60% of the rain gauge stations reported <15 mm of rainfall, whereas a few stations had high amounts up to 55 mm.

The effects of rain gauge adjustment of radar images for each step of the ARNE algorithm are shown on Fig. 5. The main land area

Fig. 4. Map of the Skjern catchment. The Skjern River network as well as observation points for precipitation and stream discharge stations are shown. Discharge stations no. 20020, 20082, and 20078 are marked for further discussion. 
of Denmark including Zealand, Fuen, and Jutland were considered, whereas the sea areas and the island of Bornholm were left out of the figures. The radars used in this example were Roemoe, Sindal, and Stevns (see Fig. 1 for locations), and the composite scenario for merging radar images was the maximum pixel values. As a primary quality control, 359 valid \( G-R \) pairs were used for adjustment and 13 pairs were rejected for representativeness reasons. The calculated decorrelation distance was found to be \( D_0 = 56 \text{ km} \).

The distance-dependent adjustment factor field \( (F) \) is shown in Fig. 5a. It is evident that the signals returned from the Roemoe radar were the strongest in that a large numbers of pixels close to the Sindal and Stevns radar sites were taken over by the signals from the Roemoe radar. Figure 5a also reveals stronger distance dependency in the areas close to the radar sites. The first-guess adjustment factor field \( \hat{F}_{(i,j)} \) shown in Fig. 5b gives much higher values in the northwestern part of Jutland, which indicates higher mean bias values between radar-estimated and rain-gauge-measured rainfall in that area. The interpolated adjustment factor field \( \hat{F}_{(i,j)} \), Fig. 5c, clearly reflects the local density of the rain gauges, and a lighter color denotes larger spatial variations (Eq. [9]). On the final adjustment factor field \( R_{(i,j)} \), Fig. 5d, the highest value reached 0.726, indicating that some pixels on the original radar image underestimated the rainfall by a factor of 5.3. This value is considered rather high and speculations have arisen that the radar hardware may have had a malfunction at some point. Ultimately,
$R_{ij}$ was applied in Eq. [11], which completed the adjustment of the raw radar image.

Following the steps above, radar precipitation estimates before and after the ARNE adjustment are shown in Fig. 6a and 6b, along with the rain-gauge-generated precipitation products in Fig. 6c and 6d. The raw radar composite image suffers from a severe underestimation of rainfall. The system bias has been largely removed in the ARNE-adjusted image and more details on the precipitation field are revealed. The DMI10 is the most widely used data source in Denmark even though it has the coarsest spatial resolution. Its spatial distribution coincides with what can be obtained from ARNE and kriging of rainfall data from the rain gauge stations. By visual inspection, the kriging method performed just as good as the inverse distance method although based on fewer observation points.

**Performance of Radar Precipitation Estimation against Ground Measurement**

A quantitative evaluation for the case study was performed. To do that, the total number of rain gauges was divided into two groups. The rain gauges that were still in operation after 2006 were used for radar image adjustment in ARNE, the removal of the mean field bias, and the spatial interpolation using kriging. The rain gauges that were closed in 2007 were regarded as reference gauges subjected to independent validation. The locations of these two groups

---

**Fig. 6.** Estimation of precipitation using different approaches for Denmark on 27 June 2006: (a) original radar image, (b) ARNE-adjusted radar quantitative precipitation estimation product, (c) Danish Meteorological Institute 10-km grid product, and (d) kriging of rain gauge rainfall data.
of rain gauges are shown in Fig. 3a and 3b. The objective functions used for the performance evaluation were bias (both absolute and logarithmic), RMSE, and the correlation coefficient ($r$).

The absolute bias between different rainfall measures and colocated validation gauges for the case study (Fig. 6) on 27 June 2006 are shown in Fig. 7a. Among them, ARNE mostly stayed bias neutral but shows overestimation under small rainfall intensity and underestimation in the range of higher rainfall intensity. This indicates that the systematic bias in radar QPE caused by raindrop size variations was not completely removed. The raw radar data show severe underestimation, which confirms the observation in Fig. 6a. The kriging result shows overestimation in the lower rainfall intensity range, which is a result of the kriging algorithm neglecting the observation points with zero precipitation. The MFB-corrected radar data largely removed the bias but was still inferior to ARNE. In summary, merging of radar and rain gauge estimation showed significant improvements in bias reduction compared with the case where information from either radar or rain gauge was used alone.

Figure 7b explicitly illustrates the relation between logarithmic bias and distance from the validation points to the radars. Without any means of adjustment, the raw radar data show a minimum bias at 50 km, which then increases with distance. This curve offers a good example of the errors residing in the process of radar QPE. At short distance, the radar suffers from partial beam blockage, whereas at longer distance, underestimation occurs when the radar detects a hydrometeor with a smaller size at higher altitude and the radar beams overshoot the main precipitating clouds. In the current case study, the meteorological condition was summer convective rain so that there was no presence of melting ice. During the wintertime, however, melting ice crystals coated with water layers will have extremely high reflectivity, which would result in a bright area on the radar display (the so-called “bright band”). Therefore, this curve may have a different shape during the wintertime with the same precipitation type when a bright band is present. The MFB corrects the radar image such that it increases or decreases the overall precipitation field as a whole. Hence, no impact on the distance relation was found except that the curve of the raw radar data was shifted downward. The ARNE successfully removed the distance-induced bias. It was also able to keep the bias considerably small and almost constant with an increase of distance.

The case study for 27 June 2006 was selected as a demonstration of daily precipitation; however, the general precipitation conditions throughout the whole year are also essential. In Table 3, the results of a performance evaluation for 2006, using all wetted validation gauges, are listed. The ARNE method was superior to the other precipitation products in every aspect. The ARNE is also the only product that had positive bias. The raw radar data provided the least promising rainfall product. These data led to the most severe underestimation of −0.79 mm/d, which accounts for >20% of the daily average precipitation for that year. Possible explanations can be inappropriate use of the $Z$–$R$ relationship, an underdeveloped clutter and noise cancellation program, problematic radar hardware calibration, or a combination of these factors. It again
emphasizes the importance of a sophisticated clutter filter and the use of rain gauge adjustment procedures applied to raw radar images given the current condition of raw data quality. The MFB was the second most favorable rainfall product after ARNE, especially in the removal of bias and the RMSE. Kriging had the worst performance in all categories and therefore was not considered for the hydrologic analysis.

Overall, radar-based rainfall estimation had a higher accuracy than rainfall values projected by rain gauge interpolation at the same locations, which clearly suggests that the combination of radar and rain gauge captures more spatial variability of precipitation than traditional rain gauge measurements. Therefore it was expected that radar QPE would influence the results of hydrologic modeling when switching from rain-gauge-based rainfall input. The ARNE and the MFB methods were the tools used to study the impact of precipitation on the hydrologic responses.

**Impact of Radar Quantitative Precipitation Estimation on Water Balance Modeling**

One of the primary objectives of this study was to assess the utility of radar QPE from a hydrologic modeling perspective. To achieve that, daily radar precipitation products were prepared using the ARNE and MFB methods for the Skjern catchment. Along with the rain-gauge-derived precipitation products, Fig. 8 illustrates four precipitation estimates that were imported to the hydrologic model in the form of the accumulated precipitation amount in 2006. It should be mentioned that accumulation of the radar-estimated precipitation inevitably resulted in magnified static ground clutter and artifacts. It was decided in this study to identify the pixels in the top 2% of the histogram and replace the erroneous data by averaging the values from adjacent pixels. This is a common methodology in the interpretation of remote sensing data and reduces the impact of data contamination on the water balance modeling.

Mean areal precipitation estimated from rain gauge observations and Thiessen polygons had the coarsest spatial resolution (Fig. 8a). It was also the rainfall product used for the calibration of the hydrologic model (Stisen et al., 2011). The combination of radar and rain gauge observations showed comprehensible advantages in that ARNE not only provided much more detail than DMI10 but also extended the area with high rainfall in the northern part of the catchment to the northeast corner. This area is interesting from a modeling perspective and is discussed below. The accumulated

![Fig. 8. Precipitation estimates for 2006 for the hydrologic model of the Skjern catchment: (a) rain gauge station measurement with Thiessen polygon, (b) Danish Meteorological Institute 10-km grid product, (c) ARNE algorithm adjusted radar quantitative precipitation estimation product, and (d) mean field bias (MFB) corrected raw radar image.](image-url)
precipitation was not expected to exhibit drastic spatial variation at the local scale, which makes the smoother transition of the rainfall field in ARNE more plausible compared with the DMI10. The MFB-removed product also showed a smoother spatial distribution than the rain gauge products. We noted that MFB indicated an entirely different rainfall distribution than the other data sets. The high-rainfall area tended to shift to the west side, which is close to the fjord. This reveals that the original radar images actually captured a considerable amount of precipitation at places where few rain gauge sampling points were located. This was not a significant issue in this study because the fjord was not accounted for in the calculation of the total water balance, but in catchments with large open water bodies surrounded by land, radar data may have more viability than a rain gauge network.

Statistics on the precipitation in the Skjern catchment during 2006 are shown in Fig. 9. To plot the frequency distribution in Fig. 9a, a bin width of 5 mm/d was chosen and equal spacing was applied except for the first band, which was 0.5 mm/d. Consistent agreement for the different precipitation products was observed in every band thanks to the bias adjustment. More than 40% of the days were either dry or had precipitation <0.5 mm/d. It was also recognized that precipitation events in the 10 to 15 mm/d interval had a more frequent occurrence than precipitation in preceding and succeeding intervals. In Fig. 9b, the temporal distribution of precipitation on a monthly basis is shown, revealing that the agreement between the radar products and the gauge product was acceptable in most cases. It also shows that the last 3 mo of the year accounted for more than half of the annual precipitation, which makes it the time of interest to observe hydrologic responses caused by precipitation forcing.

In Fig. 10, the simulated hydrographs at the three discharge stations indicated in Fig. 4 are shown. The response of runoff to precipitation was the most instant process compared with other hydrologic responses such as groundwater level variations. As radar data were available for just 1 yr, stream discharge was an ideal indicator to explain the impact of radar precipitation on the hydrologic model response. At the smaller scale represented by the upstream station (Fig. 10a), the choice of precipitation product had a significant impact on the simulated discharge. The ARNE method generated higher precipitation in the catchment draining to this discharge station (see Fig. 8c), which resulted in higher stream discharge. Especially during the event that occurred in mid-December, the ARNE input resulted in much higher peak flow than the other input scenarios. The consequences of model calibration become obvious at the smaller scale because only the baseline model (precipitation estimated by station data with Thiessen polygons) coincides with the observation during the event in December; however, the dynamics of the three hydrographs are comparable. On the other hand, all analytical rainfall products lacked the ability to reproduce many of the small peaks found in the observations. This is presumably due to the uncertainties in the hydrologic model structures rather than insufficient information on the spatial and temporal distribution of the precipitation.

In contrast to the relatively large discrepancies found between the three simulations at the upstream station, much closer agreement was observed at the two downstream stations (Fig. 10b and 10c). Simulated discharges from the radar products were able to reproduce the volume and timing of the major peak flows. It can be seen that the peak flows are closely related to the mean quantity of precipitation, which indicates that the simulated stream discharge is more sensitive to the integrated volume of rainfall entering the system rather than the spatial variation of precipitation as the catchment size increases. Therefore, radar QPE is expected to have a higher potential for improving the performance of hydrologic modeling at the small scale. The peak flow from mid-May to early June was not captured by any of the precipitation input scenarios or the baseline model. This indicates that the current hydrologic model needs more advanced model descriptions or calibration. The precipitation of the MFB method was much higher.
in the southern part of the catchment, which caused the higher discharges at Station 20078 (Fig. 10c). As mentioned above, MFB preserves the original distribution of the radar rainfall field, while ARNE reshapes the precipitation field based on the rain gauge properties. Therefore, we believe that the differences observed in Fig. 10 can be resolved by using wind correction factors for individual rain gauges on a daily basis instead of the universal monthly correction factors implemented currently. It is inconclusive, however, which rainfall distribution is more dependable at this stage.

As a primary discovery found from the simulated stream discharge, the sensitivity of the model predictions highly depended on the scale of the catchment size. This theory was further investigated by delineating the Skjern catchment into 15 subcatchments (Fig. 11), each with an actual discharge station at the outlet. The water collection areas ranged from 41 to 1553 km². The mean areal runoff was calculated by dividing the accumulated discharge volume at each outlet by the size of the subcatchments individually. Figures 12a and 12b show high dynamics at smaller subcatchment scales for the simulated discharge volume determined by all three precipitation products. The fluctuations tended to decrease with an increase in the subcatchment size. Comparing the modeled results from the radar products (ARNE and MFB) and the rain gauge product (DMI10), large differences occurred when the subcatchment size was less than around 400 km². This confirms our previous finding in a quantitative manner that the radar precipitation can effectively influence the simulated stream discharge when the catchment size is smaller than roughly 400 km² for the current Skjern River model. Nash–Sutcliffe coefficients were also calculated for the simulated daily discharge driven by radar precipitation to capture the dynamic differences between the simulations. In this process, the simulated discharge from DMI10 was assumed to represent the observed values. Figure 12c depicts that the simulations driven by radar-based precipitation deviated significantly from the DMI10 at a small scale. This is especially clear for Station 250091, where a Nash–Sutcliffe coefficient for the MFB of nearly zero was found, which is also consistent with the spatial distribution of precipitation observed from Fig. 8.

The simulated recharge to the groundwater is shown in Fig. 13. The importance of the rainfall products selected is underscored because the simulated groundwater recharge reflects exactly the same spatial pattern as the precipitation input (Fig. 8). The radar-based simulations provided a smoother and hence more realistic distribution of recharge.
compared with the two station-based simulations, where the boundaries between the precipitation zones are easily recognized. In addition, large differences in the spatial distribution of groundwater recharge were found. This is expected to result in long-term effects on groundwater levels that are not captured by the present simulations with a duration of only 1 yr.

Total water balances for the three models are summarized in Table 4 and a diagram depicting each element in the table is shown in Fig. 14. For annual precipitation, the baseline model returned the lowest value while the value estimated by the ARNE method was around 50 mm or 5% higher. The ARNE algorithm, being the candidate for a future radar QPE product in Denmark, proved to have sufficient accuracy in serving as input to distributed modeling of the seasonal behavior of hydrologic processes. The outcomes of the ARNE method were very close to that from the baseline model, which is considered promising.

**Conclusions**

A methodology for the generation of a daily radar QPE product by using the DMI ARNE algorithm was introduced. The ARNE intermediate products as well as the final product were evaluated, indicating that the radars in the network are calibrated inadequately, with an underestimation of precipitation of around 20%. The bias correction procedures in ARNE, however, successfully removed both the systematic mean bias and the distance-induced bias. The performance evaluation in 2006 for different precipitation estimators suggested that merging radar and rain gauge measurements could improve the overall data quality over using a single data source. The resulting product revealed more spatial details of the precipitation field that might not be captured by interpolation of rain gauge data. The effect of improved spatial distribution and resolution was demonstrated by simulated stream flows of the hydrologic model.

---

Fig. 11. Delineation of the Skjern River catchment to 15 subcatchments based on the existing stream discharge stations.

Fig. 12. Results from the MIKE-SHE model scaling studies comparing (a) the Danish Meteorological Institute 10-km grid product (DMI10) and ARNE algorithm adjusted radar quantitative precipitation estimation product, and (b) the DMI10 and the mean field bias (MFB) corrected raw radar image (annual mean simulated discharge is plotted against catchment size), and (c) the Nash–Sutcliffe coefficient calculated with the assumption that the simulated discharge by DMI10 is the observation.
Fig. 13. Simulated groundwater recharge from the MIKE-SHE model: (a) result using precipitation input with station data and Thiessen polygons, (b) result using precipitation input with the Danish Meteorological Institute 10-km grid product (DMI10), (c) result using precipitation input with ARNE-adjusted radar quantitative precipitation estimation, and (d) result by using precipitation input with the mean field bias (MFB) corrected raw radar image.

Table 4. Simulated total water balance for the Skjern River catchment using scenarios of baseline precipitation, the Danish Meteorological Institute 10-km grid product (DMI10), the ARNE algorithm adjusted radar quantitative precipitation estimation product and mean field bias (MFB) corrected raw radar data as precipitation inputs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>DMI10</th>
<th>ARNE</th>
<th>MFB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>1054</td>
<td>1066</td>
<td>1117</td>
<td>1103</td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>588</td>
<td>592</td>
<td>599</td>
<td>587</td>
</tr>
<tr>
<td>Subsurface storage change</td>
<td>117</td>
<td>117</td>
<td>136</td>
<td>133</td>
</tr>
<tr>
<td>Overland storage change</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Overland flow to river</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>Base flow to river</td>
<td>186</td>
<td>187</td>
<td>195</td>
<td>191</td>
</tr>
<tr>
<td>Groundwater drain to river</td>
<td>80</td>
<td>84</td>
<td>89</td>
<td>90</td>
</tr>
<tr>
<td>Irrigation</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Pumping</td>
<td>17</td>
<td>17</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Error</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 14. Hydrologic fluxes simulated by the hydrologic model.
The main contribution of this study is the application of radar QPE for long-term water resources modeling. After having demonstrated that radar gives promising results in estimating precipitation from a meteorological perspective, radar-based daily precipitation products were brought into a distributed water resources model to analyze the hydrologic responses. Simulated stream discharge with the radar QPE as model input gave reliable results with reference to what was obtained from the baseline model. The hydrologic model also showed high sensitivity to the spatial distribution in relatively small subcatchments with sizes of <400 km², while for larger areas, it was more dependent on the total precipitation volume. The importance of using the radar QPE in hydrologic modeling is also illustrated by the close match between the simulated annual groundwater recharge and the precipitation products used as model input. The differences of the precipitation components between the three water balance model scenarios were found to be <5%. In addition, the study revealed some model development issues such as model parameter calibration.

The most urgent task for the future is to improve data quality in radar measurement. This will include a refined hardware calibration strategy and a clutter removal program. Longer time series of radar data are also needed for more thorough testing and to facilitate calibration of the hydrologic model using radar data as input. We also plan to operate radar QPE on a subdaily basis, not only to meet the requirement for more precise hydrologic modeling but also to accommodate automation of the rain gauge network.

References


STATISTICAL ANALYSIS OF THE IMPACT OF RADAR RAINFALL UNCERTAINTIES ON WATER RESOURCES MODELING

Xin He, Jens Christian Refsgaard, Torben O. Sonnenborg, Flemming Vejen, Karsten H. Jensen

Submitted to Water Resources Research, March 2011
Minor revision requested by the Associate Editor, June, 2011
Revised manuscript submitted, July 2011
**Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling**

Xin He\(^1,3,*\), Jens Christian Refsgaard\(^2\), Torben O. Sonnenborg\(^2\), Flemming Vejen\(^3\), Karsten H. Jensen\(^1\)

\(^1\) Department of Geography and Geology, University of Copenhagen
\(^2\) Department of Hydrology, Geological Survey of Denmark and Greenland
\(^3\) Data and Climate Division, Danish Meteorological Institute

* Corresponding Author (xh@geo.ku.dk)

**Abstract**

Uncertainty analysis in hydrological modeling has become an essential step in the scientific interpretation of model results and a useful tool to support decision making. Among many uncertainty sources in the modeling practice, uncertainties in precipitation estimation play an important role since it is the main driving force for other hydrological processes. The present study demonstrates a statistical method for generating radar rainfall realizations that account for the uncertainties in radar based Quantitative Precipitation Estimation (QPE). The random sampling technique used to generate stochastic uncertainty fields is based on Sequential Gaussian Simulation. The hydrological impact of the uncertainties in radar QPE is analyzed by propagating the rainfall ensemble through a distributed and integrated water resources model. The study shows that the uncertainty of the simulated stream discharge depends on the intensity of the rainfall input signal. The coefficient of variation is calculated for simulated stream discharge and groundwater recharge at subcatchments with various sizes. The results reveal strong scale dependency showing higher variations of hydrological uncertainties at smaller catchments, especially for catchment areas smaller than 50 km\(^2\). The uncertainties from precipitation input are generally amplified in the hydrological model. This effect is less obvious for groundwater recharge but rather substantial for stream discharge, where the coefficient of variation increases by a factor of three.

**Key words:**
Precipitation, weather radar, hydrological modeling, uncertainty analysis, rainfall ensemble
1. Introduction

The importance of uncertainty analysis in hydrological modeling has gained increasing attention in recent years to support decision making in water management [Beven and Binley 1992; Butts et al. 2004; Matott et al. 2009]. The key sources of uncertainty in hydrological model predictions are related to the model formulation, the parameter values, and the model inputs [Refsgaard et al. 2007]. In a recent study, Refsgaard et al. [2010] also pointed out that the future challenge in integrated water resources modeling is not only to match the output of the model to observation data, but also to inspect whether the predictions are meaningful through quantitative uncertainty analysis. This goal is difficult to achieve because the model input uncertainties are difficult to assess for the hydrological modelers.

Precipitation represents a key input to hydrological models and uncertainties in precipitation measurement account for a large fraction of the total input uncertainties in hydrological modeling. For instance, in a simple rainfall-runoff model that uses Horton runoff mechanism, precipitation is the only model input and therefore the most important factor determining the simulated stream flow [Kuczera and Williams 1992; Sun et al. 2000; Borga 2002; Smith et al. 2004; Morin et al. 2009]. Precipitation also plays an important role in the simulation of water balance and other hydrological behaviors at catchment and subcatchment scale where the spatial variability of rainfall can significantly influence the simulated catchment responses [He et al. 2011].

During the past half century weather radar has gradually been accepted as a valuable instrument for precipitation estimation [Marshall et al. 1947; Collier 1989; Rinehart 1997]. Meanwhile, Quantitative Precipitation Estimation by weather radar (radar QPE) has become a typical rainfall product of meteorological offices across Europe and North America [Wilson and Brandes 1979; Fulton et al. 1998; Harrison et al. 2000; Tabary 2007]. In principle, weather radar enables three-dimensional observation of the atmosphere with good areal coverage and high spatial/temporal resolution compared with the traditional rain gauge based rainfall products, which makes it an ideal tool for hydrological applications.

Hydrological studies employing weather radar have shown promising results in simulations of stream discharge, especially in flash flood forecasting [Finnerty et al. 1997; Carpenter et al. 2001; Cole and Moore 2008, 2009]. However, radar QPE suffers from many sources of uncertainties primarily because radar measures precipitation remotely and indirectly. Therefore, rain gauge measured rainfall is still considered as the ground truth of precipitation at point scale in most studies [Austin 1987; Joss and Lee 1995; Smyth and Illingworth 1998b]. The origin of radar QPE uncertainties may come from the radar hardware, e.g., radar miscalibration; the radar signal processing, e.g. attenuation and anomalous propagation correction; or the QPE algorithm, e.g. inappropriate use of Z-R relationship [Kitchen and Jackson 1993; Pamment and Conway 1998; Smyth and Illingworth 1998a; Anagnostou et al. 1999; Holleman 2007]. A comprehensive review of the possible radar QPE uncertainties caused by instrumental imperfections can be found in Villarini and Krajewski [2010]. The uncertainties in radar QPE can be largely reduced with advancing hardware and software technologies, but they can never be entirely eliminated.
Knowing the sources of uncertainties in radar QPE, it is essential that the uncertainty characteristics are analyzed and quantified. The goal is not only to provide the end users of the QPE products with the radar estimated rainfall but also to provide an idea of how this estimation may correspond to the actual rainfall. Previous studies have shown that many of the radar uncertainties can be formulated based on physical considerations [Bech et al. 2000; Vignal et al. 2000; Anagnostou et al. 2004; Brandes et al. 2004]. In such processes, each uncertainty source is targeted individually, and the accuracy of the correction will depend on the availability of both knowledge and specific information concerning the error sources. However, in hydrological studies that employ weather radar, it is more common that radar precipitation is considered as one component, and it is therefore more relevant to view the various radar uncertainties as a whole. Besides, separating different uncertainty sources in QPE for hydrological application is neither necessary nor technically feasible.

In recent years, a few studies have attempted to characterize the combined effect of total radar rainfall uncertainties using a statistical approach that includes the bias, the spatial-temporal dependency and the random errors altogether [Borga et al. 2006; Ciach et al. 2007; Habib et al. 2008; Germann et al. 2009; Villarini and Krajewski 2009]. As a result, the uncertainties in radar QPE are expressed by an ensemble of precipitation fields with each ensemble member representing a possible realization for a given radar rainfall measurement using the knowledge of the radar error structure. Another advantage of using the statistical approach is that the ensemble members can be used directly as input information to the hydrological models, so that quantification of the uncertainty propagation through the modeling process can be more convenient. However, most studies using radar uncertainty ensembles for hydrological applications are limited to relatively short term stream discharge simulation or flood forecasting. The impact of radar precipitation uncertainties on other hydrological processes over longer periods has not been thoroughly investigated. Moreover, many of the existing methodologies are highly demanding on computational power, which limits the suitability for studies of long term water resources modeling. Last but not least, the scaling issue in such uncertainty analysis, namely the relationship between the model input uncertainty and the model responses at various catchment sizes, needs to be further addressed.

The objectives of this study are: (1) to develop a statistical method to quantify uncertainties in radar rainfall estimation by using a less computationally expensive stochastic ensemble generator; (2) to propagate the uncertainty in radar QPE through an integrated hydrological model; and (3) to analyze the hydrological model simulation uncertainty in response to model input uncertainties at different catchment scales.
2. Method

2.1 The radar based QPE algorithm

The radar QPE algorithm used in the current study is a range-dependent accumulated precipitation estimation algorithm. The method is developed based on the protocols described in the Baltic Sea Experiment (BALTEX) data center report [Michelson 2000]. The current QPE product, short name ARNE, is a countrywide daily precipitation estimation product projected on 2 km grids. The rainfall intensity to radar reflectivity relationship is based on the classical Marshall-Palmer equation with standard parameter values $A = 220$ and $b = 1.6$ [Marshall et al. 1947]. The bias between the radar estimate and the measurements is characterized by the logarithmic gauge-to-radar ratio at daily time step, $F_g$

$$F_g = \log \left( \frac{G}{R} \right)$$  \hspace{1cm} (1)

where $G$ is the rain gauge observation values and $R$ represents the radar pixel values at the locations where the rain gauges are located. Based on the available $G$-$R$ pairs, an adjustment factor field depicting the relationship between the distance from the radars and the measurement bias can be formulated. This is done by using a second order polynomial to reflect the nonlinear relationship between the systematic variations of radar reflectivity and the distance,

$$F_r = a + br + cr^2$$  \hspace{1cm} (2)

where $r$ is the distance to the radar and $a$, $b$, $c$ are dimensionless parameters. After fine-tuning the adjustment factor ($F_r$) by using several spatial averaging techniques [Michelson 2000], a bias adjustment factor is calculated for every pixel on the radar composite image. The final radar composite sum ($R$) is obtained by multiplying the adjustment factor field ($\overline{F}$) to the original radar composite sum ($R_0$),

$$R = R_0 \cdot 10^{\overline{F}}$$  \hspace{1cm} (3)

It needs to mention that $F_g$ consists of the radar-rain gauge data pairs, and is therefore a discrete variable. $F_r$, on the other hand, is derived from $F_g$ using the polynomial fitting between the values of the $G$-$R$ pairs and the distances to the corresponding radars, and is therefore a continues variable. $\overline{F}$ is further developed based on $F_r$ which accounts for not only the distance but also the rain gauge observation density. It can be seen that ARNE is developed on the basis of the relative rainfall spatial distribution measured by the radar and merged with the more trustworthy point measurements from the rain gauges. A detailed description of the implementation of the BALTEX method can be found in [He et al. 2011].

2.2 The ensemble generator

The following describes a methodology developed in this study for generating probabilistic radar QPE ensembles which subsequently are used in the uncertainty analysis of hydrological model response to radar precipitation input uncertainties. The probabilistic radar rainfall ensemble
generator is designed to produce realizations for the original radar QPE at each time step based on the estimated error characteristics both in space and time.

The basic concept of this method is obtained from Germann et al. [2009], while the implementation is modified from the original version in order to address the problems in the present study. In the method proposed by Germann et al. [2009] the precipitation field for each ensemble member is obtained by superimposing a stochastic perturbation field on the unperturbed radar precipitation field. This concept can be expressed as,

\[ \Phi_{t,i} = R_t + \delta_{t,i} \]  

where \( \Phi \) is the precipitation field for an ensemble number representing the probabilistic term; \( R \) is the unperturbed radar precipitation representing the deterministic term; \( \delta \) is the stochastic perturbation field; \( t \) is the time step; and \( i \) is the ensemble number.

Since most radar errors are multiplicative, the residual error (\( \varepsilon \)) of radar QPE can be obtained by comparing to the true precipitation and express it in the logarithmic form,

\[ \varepsilon_t = \log(S_t / R_t) \]  

where \( S_t \) is the hypothetical true precipitation field at ground level, which is unknown under most circumstances. It needs to mention that in Eq. (5), \( \delta \) is replaced with \( \varepsilon \). That is because \( \varepsilon \) characterizes the error model in the radar QPE as a whole; therefore it is the population, whereas \( \delta \) is one realization randomly drawn from \( \varepsilon \) therefore denotes the sample. When the volume of \( \delta \) is large enough, it can sufficiently represent \( \varepsilon \). The expression of Eq. (5) is also in line with what has been defined in the ARNE method for error characterization, see Eq. (1). The stochastic term of Eq. (4) can be interpreted as one realization of the residual error in the error space. A sufficient number of realizations, as an ensemble, is required to represent the features of the residual error, however, this number may be difficult to identify. Using log transformed terms, Eq. (4) then becomes,

\[ \log[\Phi_{t,i}] = \log[R_t] + \delta_{t,i} \]  

where \( \delta_{t,i} \) now represents a perturbation of the log-transform field. Having the basic concept of the ensemble generator established, the next task is to identify the stochastic term of Eq. (6) using the uncertainties in radar QPE. First, it is assumed that the residual error of radar QPE is normally distributed, and at any point in space the error structure can be formulated as,

\[ \varepsilon_t = N(\mu_t, C_t) \]  

where \( N \) is the Gaussian random vector, \( \mu \) is the mean error, \( C \) is covariance matrix. To determine the error structure in practice, the approach is to make comparisons between radar QPE and ground reference given by the rain gauge measurements. This approach is widely acknowledged by many other studies [Woodley et al. 1975; Fulton 1999; Rutgersson et al. 2001; Ryzhkov et al. 2005; Rubel and Brugger 2009]. However, even with the most advanced product,
a rain gauge is still far from a perfect instrument. In fact, uncertainties from the rain gauge measurement are hardly to be neglected.

Let $x_k$ be the location of a radar pixel where a rain gauge is collocated, where $k$ is within $1 \cdots M$ and $M$ is the number of pixels in the image. Then Eq. (5) is rewritten as,

$$\hat{\varepsilon}_{t,x_k} = \log\left(\frac{G_{t,x_k}}{R_{t,x_k}}\right)$$

where $S$ is replaced by the rain gauge observation $G$, and $\hat{\varepsilon}$ is an estimate of the true error. If the rainfall measured at time step $t$ and location $x_k$ is temporally correlated with observations at the same location from previous $Q$ time steps, the expected mean error, $\hat{\mu}_{x_k}$, can be calculated correspondingly,

$$\hat{\mu}_{x_k} = \sum_{i=1}^{Q} r_{t-1,x_k} \cdot \omega_{t-1,x_k} \cdot \hat{\varepsilon}_{t,x_k} \bigg/ \sum_{i=1}^{Q} r_{t-1,x_k} \cdot \omega_{t-1,x_k}$$

where $Q$ is the number of time steps, $\omega_{t-1,x_k}$ is the weight of observed precipitation where we force the value to be equal to the original unperturbed radar pixel value. $r_{t,x_k}$ is the autocorrelation coefficient defined as,

$$r_{t,x_k} = \frac{\sum_{i=1}^{Q} (R_{x_k,t} - \bar{R}) \cdot (R_{x_k,t+i} - \bar{R})}{\sum_{i=1}^{Q} (R_{x_k,t} - \bar{R})^2}$$

where $\tau$ is the time lag, $R_{x_k,t}$ is the rainfall measured at location $x_k$ at time step $t$, $Q$ is the total number of time steps which are believed to be temporally correlated, and $\bar{R}$ is the averaged rainfall over $Q$ time steps. For clarification, $r_0$ indicates the autocorrelation coefficient by itself and $r_1$ is for one time step apart, etc. It is assumed that the covariance of the error structure in space and time are independent; therefore it has the additive feature,

$$\hat{C}_{x_k,x_l} = \frac{(\hat{\mu}_{x_k} - \hat{\mu}_{x_l})^2}{2} + \sum_{i=1}^{Q} r_{t-1,x_k} \cdot r_{t-1,x_l} \cdot \omega_{t-1,x_k} \cdot \omega_{t-1,x_l} \cdot (\hat{\varepsilon}_{t,x_k} - \hat{\mu}_{x_k}) \cdot (\hat{\varepsilon}_{t,x_l} - \hat{\mu}_{x_l})$$

$\hat{C}_{x_k,x_l}$ is the expected covariance between radar pixel $x_k$ and $x_l$, for $Q$ time steps. The first term on the equation indicates the error covariance in space, and the second term represents the temporal correlation among $Q$ time steps. The subscripts are the same as explained previously.

Eq. (9) and (11) provide the estimated error mean and covariance at locations where rain gauges exist. For the rest of the grid cells in the radar domain, the values need to be interpolated based on the values from the gauged locations and at the same time incorporate the idea of stochastic
sampling. Therefore, ordinary Kriging coupled with Sequential Gaussian Simulation (SGS) is used. Kriging is one of the most widely use interpolation methods in geosciences. SGS is applicable since it has already been assumed that the error structure follows the Gaussian distribution, see Eq. (7).

The SGS approach uses a random visiting path to the model grid nodes. Each new simulated value is conditioned on the observation data and previously simulated values, where ordinary Kriging is used to obtain the mean and covariance. The estimate of the random variable \( \varepsilon \) at a point in space is calculated by randomly drawing values from the normal distribution error model using the kriging mean and variance. Therefore, the implementation of SGS involves two sources of random processes: a random grid visiting and a random Gaussian sampling. The random grid visiting implies that the path is changed from realization to realization. In the original description of the SGS, no assumptions were made with respect to the order the unsampled grid nodes are visited. Since previously simulated values affect the conditioning of the subsequent simulation, random starting point and random path are used in this study to prevent artifacts or artificial correlations.

Note that the quantity being interpolated and sampled in this process is the error of radar QPE and not the QPE itself. In addition to radar and rain gauge measurements, SGS also requests a variogram model as one of the input terms. Prior to this study, a linear model and an exponential model were tested for a few case studies, and they performed equally well mainly because the precipitation in the research area is rather homogeneous and the correlation distance is long. In the present study, an exponential model is used to reflect the relationship between the semivariance and the distance since it is known that the exponential model is superior to the linear model in defining the correlation length in case of a local rainfall condition,

\[
\hat{\gamma}(h) = c \cdot [1 - \exp(-\frac{h}{a})] + c_0
\]  \hspace{1cm} (12)

where \( h \) is distance to the observation points, and \( a \) (range), \( c \) (sill), \( c_0 \) (nugget) are the readings from the experimental semivariogram. With the above procedures completed, the stochastic term of Eq. (6) is obtained and the final form of the ensemble generator is shown below,

\[
\Phi_{t,i} = R_{t} \cdot 10^{\delta_{t,i}}
\]  \hspace{1cm} (13)

This expression is consistent with the paradigm used when the original radar composite sum was adjusted, see Eq. (3).

2.3 Study area

The Skjern catchment is the largest river basin in Denmark in terms of stream discharge volume. This makes the Skjern catchment one of the most delicate areas in Denmark for water resources exploitation and protection considering the more intensified climate changes in the future [van Roosmalen et al. 2009]. The geographical location of the catchment relative to the whole country is seen on Fig. 1 and a close up on Fig. 2. The catchment covers an area around 3500 km\(^2\).
excluding the Ringkoebing Fjord. Land use consists of 86% agriculture, 5% heath, 7% forest and 2% urban areas [van Roosmalen et al. 2007].

The climate regime over the Skjern catchment is typical maritime climate. The dominant westerly wind is accompanied by mild winters and relatively cold summers with frequent rainfall events all year round. The mean annual precipitation is estimated to 1057 mm and the mean annual reference evapotranspiration is 575 mm. Precipitation measured at 89 rain gauges in the catchment and the vicinity (depicted on Fig. 2) over 5 years are shown on Fig. 3. Maximum precipitation occurs in autumn and minimum in spring.

As seen on Fig. 2, land surface elevation gently slopes towards west and decrease from 125 meters to sea level at the coast. The Skjern River and its tributaries follow the glacial outwash plains running from east to west. The mean annual stream discharge is about 480 mm. The geology in the research area is mainly composed of sand and gravel which forms large interconnected aquifers [van Roosmalen et al. 2007]. Because of the highly permeable soils and geological settings the discharge to the streams in the Skjern catchment is dominated by groundwater flow.

![Illustration of radar coverage over the Skjern catchment area by the DMI operational radar network. Radar detection range is shown in 240 km. Roemoe, Sindal and Stevns radars are each color coded correspondingly.](image)

**Figure 1.** Illustration of radar coverage over the Skjern catchment area by the DMI operational radar network. Radar detection range is shown in 240 km. Roemoe, Sindal and Stevns radars are each color coded correspondingly.
**Figure 2.** Locations of rain gauges in HOBE catchment. Topography is shown in the background. Gauges used in ARNE (Group 1) and QPE uncertainty analysis (Group 2) are marked respectively. Discharge stations 250082, 250020 and 250021; rain gauge station 24330 are marked for discussion.

**Figure 3.** Monthly rainfall in Skjern catchment area 2002-2006. Error bars indicate the standard deviation among the rain gauge observed values.
### Table 1. Operational parameter settings for the Danish national radar network

<table>
<thead>
<tr>
<th>Radar parameters</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>5 cm (C-band)</td>
</tr>
<tr>
<td>Range</td>
<td>240 km</td>
</tr>
<tr>
<td>Pulse length</td>
<td>2 µs</td>
</tr>
<tr>
<td>Pulse repetition frequency (PRF)</td>
<td>250 Hz</td>
</tr>
<tr>
<td>Peak power</td>
<td>250 KW</td>
</tr>
<tr>
<td>Rotation rate</td>
<td>20 °/s</td>
</tr>
<tr>
<td>Antenna gain</td>
<td>45 dB</td>
</tr>
<tr>
<td>Vertical beam width</td>
<td>± 0.5 °</td>
</tr>
<tr>
<td>Horizontal beam width</td>
<td>1.0 °</td>
</tr>
<tr>
<td>Volume scan angles</td>
<td>0.5, 0.7, 1.0, 1.5, 2.4,</td>
</tr>
<tr>
<td>Scans</td>
<td>4.5, 8.5, 13.0, 15.0</td>
</tr>
<tr>
<td>Scans</td>
<td>PPI</td>
</tr>
</tbody>
</table>

#### 2.4 Radar data

The Danish Meteorological Institute (DMI) operates five C-band Doppler radars, three of which have full or partial coverage over the area, see Fig. 1. It is seen that all radars are relatively far from the catchment considered. Roemoe radar is the closest to the catchment but still 100-120 km away. All radars are manufactured by Electronic Enterprise Corporation (EEC) with operational parameter configurations as summarized in Table 1. The visualization of radar reflectivity signals is an integration of volume scans at 2 km altitude known as a product called pseudo-CAPPI (Constant Altitude Plan Position Indicator). Based on past experiences, radars in Denmark are calibrated unequally at different locations. As a result, a boundary line can be seen at the places where radar beams are intercepted. In order to minimize this problem, the maximum output pixel value is used as the composite strategy in areas where radar detections are overlapped.

In order to implement the radar QPE algorithm and the above introduced ensemble generator, the total number of rain gauges over the Skjern catchment area is divided into two groups, see Fig. 2. Group 1 is used for raw radar image adjustment (ARNE), and Group 2 is subjected to uncertainty analysis. One reason for making this arrangement is that the rain gauges in Group 2 have been closed after 2007, while Group 1 continues to be in operation up to present day.
2.5 Hydrological model

A water resources model is developed for the Skjern catchment with the objective to explore and quantify the entire water cycle. The model is built using the MIKE SHE code which is a deterministic, distributed and physically-based hydrological modeling system [Abbott et al. 1986; Henriksen et al. 2003]. The hydrological model for the Skjern catchment is based on the Danish National Water Resources Model, the DK-Model, which has been further modified by Stisen et al. [2011]. In the present application the following model descriptions are applied: 1) the Kristensen and Jensen model [Kristensen and Jensen 1975] for evapotranspiration, 2) 2D diffusive wave approximation of the Saint Venant equations for overland flow, 3) Muskingum-Cunge routing [Chow et al. 1988] for flow in the upstream part of the river system and the full Saint Venant equations for the downstream part of the river system, 4) a two-layer water balance method [Yan and Smith 1994] for describing the distribution of infiltration between evapotranspiration and groundwater recharge in the unsaturated zone, 5) linear reservoir model for flow in drains, 6) 3D Boussinesq equation for flow in the saturated zone, 7) Darcy equation for the river–aquifer interaction, and 8) a degree-day approach for snowmelt. More information about the MIKE SHE modeling system can be found in [Refsgaard and Storm 1995; Graham and Butts 2006]. The baseline Skjern river model is calibrated using the PEST optimization tool against observed stream discharge and hydraulic head data [Doherty, 2004]. When calibrating the model, precipitation input is defined by rain gauge measurements and spatially estimated by Thiessen polygons [Stisen et al. 2011]. For the uncertainty analysis the model is run from 2002 to 2006. In each simulation, precipitation input in 2006 is replaced by simulated random rainfall realizations.
3. Results and discussion

3.1 Sampling of radar rainfall estimation error using Sequential Gaussian Simulation (SGS)

In order to perform the SGS and quantify the uncertainties in radar rainfall estimation, several assumptions have been made. One of the essential assumptions is that the probability distribution of the estimation error at each grid cell is normally distributed in space and time. This assumption not only provides a prerequisite for the implementation of SGS but also infers that the uncertainty being quantified in the long term mainly comes from the rainfall spatial and temporal variability, which is unsuccessfully captured by the instrument, not the systematic bias between the estimated rainfall and the ‘true’ rainfall. To verify this assumption, a radar pixel containing a rain gauge station, st. no. 24330 (Fig. 2), is chosen for demonstration. Among the 31 radar – rain gauge pairs inside the study domain, this location is chosen randomly. Results from the rest of the data pairs (not shown here) all follow the same pattern.

Throughout the year 2006, st. 24330 has witnessed 184 days with daily precipitation larger than zero. Based on Eq. 1 the radar rainfall estimation error is obtained and a histogram showing the error frequency distribution is established, see Fig. 4a. The peak of the histogram is slightly displaced towards the negative side indicating overestimation. This is because the radar QPE procedure, ARNE, is based on a country wide bias adjustment using all rain gauge observation data in the nation. Therefore, a complete bias removal at local scale is hard to obtain. To overcome this problem, Mean Field Bias (MFB) correction is carried out using only the local rain gauge stations (Group 1 in Fig. 2). MFB is defined as:

\[ MFB = \frac{\sum_{i=1}^{n} G_i}{\sum_{i=1}^{n} R_i} \]

where \( G_i \) is the rain gauge observations and \( R_i \) is the radar estimated values at the pixels that contain the rain gauges. Therefore, a single correction factor is multiplied to the precipitation field at Skjern catchment on a daily basis as an additive to ARNE. It is acknowledged that MFB correction is a common technique in radar meteorology to remove systematic bias [Smith and Krajewski 1991; Seo et al. 2000; Holleman 2007]. A histogram established using MFB corrected radar pixel values is shown on Fig. 4b, where the shape of the histogram fits the normal distribution more properly. A statistical test [Anderson and Darling 1952] was carried out to examine the normality of the two data sets for zero mean Gaussian distribution, where the test scores were calculated to be 2.446 before and 0.746 after the MFB correction. The 5% threshold level of significance is 0.752, which indicates the data set after the MFB correction may be normally distributed. Although it is not likely that every \( \log(G/R) \) data series passed the test, the test scores after MFB correction were always significantly better than that without the corrections; hence adding the MFB correction is preferable.

Identifying the spatial and temporal correlation in the error structure model is one of the important steps in the proposed methodology. The temporal correlation in ground measured point precipitation is typically within the range of several minutes to few hours [Gupta and Waymire 1993; Nelson et al. 2005; Lee and Oh 2006]. In the current implementation, ARNE
provides rainfall estimation at a daily basis. Hence, if the temporal correlation scale is smaller
the necessity of including the temporal term in Eq. 9 and 11 is questioned. Discarding the term
for temporal correlation while retaining the term for spatial correlation would largely save
computational time.

A dataset of all available rain gauge data from Skjern catchment during 1990 to 2006, in total 17
year, is analyzed. The autocorrelation coefficient at each rain gauge station is calculated using
Eq.10, then averaged catchment wise. While calculating the autocorrelation coefficients, two
scenarios are considered. In the first scenario, the observed rainfall values are regarded as a
source of non-interruptive signals (referred to as ‘continuous’ in the following), so that the
temporal autocorrelation is obtained by using every measurement in the dataset including zero
precipitation. In the second scenario, the rainfall record is divided into individual events (referred
to as ‘intermittent’ in the following). The ‘intermittent’ precipitation is defined as periods where
non-zero rainfall measurements are observed for five consecutive days, which eliminates the
situations with zero precipitation from the calculation.

\[ \text{Figure 4. Histograms of logarithmic bias between daily values of rain gauge measured and radar}
\text{estimated precipitation in 2006 for station 24330. Histograms are shown before (a) and after (b)
mean field bias corrections.} \]
Figure 5. Autocorrelogram of precipitation over 17 years in Skjern catchment. Two scenarios are made to consider rainfall as continuous or intermittent signals. Temporal correlation coefficients are calculated for individual rain gauge stations and then averaged.

The results calculated for both scenarios are presented in Fig. 5. Lag-0 indicates the temporal correlation of the precipitation measurement to itself, thus remains at 1 for the two scenarios. The autocorrelogram of the ‘continuous’ rainfall gently decreases to near zero after 3 time lags (3 days). However, the scenario of the ‘intermittent’ precipitation drops sharply to below zero in lag-1 and stays negative. It suggests that the temporal correlation of daily precipitation measured at Skjern catchment is characterized by white noise only, namely the zero precipitation data. Therefore, it can be concluded that when radar QPE uncertainties are assessed on the basis of daily precipitation values, it is reasonable to assume that the point measurement data are not temporally correlated. Another critical assumption in the methodology for generating uncertainty realizations is that the covariance structure is independent in space and time. This assumption is valid in the present study also based on the fact that daily time steps are used. However, it should be emphasized that this simplification has only been tested in this particular study. When applying the method under other circumstances such as using hourly precipitation data instead of daily data, catchments at different geographical locations, or different rainfall instrumentation, these assumptions need to be re-evaluated.

The main objective of the proposed methodology is to generate stochastic fields using the established error structure. Following the steps explained previously, this goal is accomplished by SGS and the results are illustrated in Fig. 6. Note that Fig. 6 only demonstrates one realization of SGS for a selected day, whereas in the rainfall ensemble much more realizations are required. The date chosen for demonstration is June 27, 2006. It represents a typical 24-hour cumulative rainfall pattern in Denmark during summer time, which is widespread rain of moderate intensity. The meteorological report by DMI shows that 60% of the rain gauges in Skjern catchment observed less than 15 mm of rainfall and the highest measurement was 55 mm.
During the SGS process the grid cells on the domain are filled successively in random order, see Fig. 6a, b, c. The final product in Fig. 6d shows a relatively high degree of randomness. To interpret the color codes, blue indicates positive logarithmic bias value which by definition corresponds to underestimation of the rainfall by radar, while red depicts overestimation. The sampling process does not fix the rain gauge coordinates at the center of each grid cell and therefore does not restrict the radar pixel values to be exactly the same as the rain gauge measurement values. However, due to the spatially correlated error field the pixels with high uncertainty values only appear in the pixels where rain gauge data is unavailable. It shows that the SGS algorithm conditions the radar rainfall estimation uncertainties based on the nearby rain gauge information, since the spatial error structure is built entirely on distance relations. The numerical range of the uncertainty field is obtained by converting the log-transformed bias back to rain gauge to radar ratio which varies from 0.74 to 1.35 for the case study (1 being bias neutral) and multiply this factor with the rain gauge values.

Figure 6. Example of Sequential Gaussian Simulation (SGS) of log-transformed radar estimation error for a selected day. The image grid cells are gradually filled in random order, the fullness/emptiness is shown after (a) 100 pixels, (b) 500 pixels, (c) 1000 pixels. (d) is the final product of SGS, the legend on (d) represents the logarithmic rain gauge to radar ratio, where positive values indicate underestimation of rainfall. The position of Skjern catchment is also overlaid. Rain gauges used in ARNE are shown in green circles, rain gauges used in the uncertainty analysis are shown in yellow circles.
3.2 Generating rainfall realizations as input to the hydrological model

Based on the uncertainty distribution of accumulated daily radar precipitation for each grid cell, the rainfall ensemble members used as input to the hydrological model are obtained by adding the stochastic error field to the unperturbed radar QPE field. Two days are chosen to demonstrate the evolution of the images patterns, June 27, 2006 and August 18, 2006. The raw radar composite images, the radar images adjusted by the ARNE algorithm and the perturbed radar images after adding the stochastic error field for the two case days are shown on Fig 7. The two days represent two completely different precipitation regimes: widespread stratiform rain and local convective rain, respectively. The situation in June exhibits higher cumulative precipitation in the northeast of the catchment, Fig. 7b, which gradually decreases to around 4 mm/day in the southwest. The August case, Fig. 7e, shows three distinctive convective cells with rainfall intensity up to 50 mm/day which each occupies a corner of the catchment leaving the rest of the area nearly dry. It is seen that the rain gauge adjustment preformed by ARNE does not only remove the bias from the raw radar image but also modifies the spatial distribution of rainfall. This effect is more evident in the August case. The radar tends to underestimate precipitation under stratiform conditions at the range used in the present study. This is characterized by beam partial filling or complete overshooting since the clouds are very low, and also maybe by small raindrop size and low reflected electromagnetic power. This effect is visually confirmed by Fig.7a and Fig.7b. Note that Fig.7c is a result of multiplying Fig. 6d onto Fig.7b. The overall perturbed rainfall fields appear to be more pixelated in the convective case than in the stratiform case because the spatial correlation length is much larger in the stratiform case.

In principle, the proposed method is able to generate an infinite number of error maps, such as in Fig. 6d, and thus an infinite number of rainfall realizations with the same likelihood. It is also recognized that the larger the ensemble size, the better it represents the uncertainty in the radar rainfall estimation without missing any critical information. The selection of the number of realizations depends on the purpose of the specific study and consequently the sampling methodology applied: some studies choose hundreds [Jordan et al. 2003; Carpenter and Georgakakos 2004] while others use tens of thousands especially when the GLUE methodology is involved [Beven and Freer 2001; Borga et al. 2006; Younger et al. 2009]. We want to keep the number of simulations to a minimum in this study due to the heavy computational demands of the MIKE SHE model.
Figure 7. Two case studies of stochastic modeling based on radar precipitation estimation on (a) – (c) June 27, 2006 and (d) – (f) August 18, 2006. Examples shown are radar images at different stages: (a, d) Raw radar composite image; (b, e) Image adjusted by ARNE algorithm, and (c, f) Examples of realizations after adding the perturbation field generated by SGS. (Unit: mm/day).
It is reassured by using the MFB correction to the ARNE rainfall field that the stochastic error field is based on a zero mean Gaussian distribution. Therefore, the average of a large amount of rainfall realizations will result in a rainfall field that is identical to the unperturbed radar QPE field. Our approach is to look for the number of realizations at which the main statistical features of the rainfall ensemble are sufficiently close to the original radar QPE. To implement this concept, first the cumulative density functions (CDFs) for the above introduced two case studies are drawn, as seen in Fig. 8. Then the 2.5%, 50% and 97.5% quantiles are retrieved from the CDFs. Note that the shapes of the two CDF curves are distinctively different due to the rainfall regimes.

Subsequently, values of precipitation taken at the same quantiles from each rainfall ensemble member is successively averaged, Fig. 9, from 1 up to 500 realizations. Initially, the results for each panel are rather random; however they all converge to the same values as the radar QPE as expected. Nevertheless, the speed of their convergence is very different at each quantile, where the lower quantile reach a constant value relatively fast. Therefore, the selection of the ensemble member is dependent on the performance of the highest quantile. Based on Fig. 9, it is reasonable to argue that increasing the number of realizations beyond 200 does not impose significant impact for the two cases. The same overall conclusion is reached for other days in 2006. It is also noticed that for the highest quantile the convective system observed in August 2006 shows more fluctuation at low ensemble size than the stratiform case observed in June. However, the overall change on daily precipitation beyond 200 realizations is lower than 1% which is considered to be a sufficient accuracy to serve the purpose of the present study.

Figure 8. Cumulative Density Functions of two case studies (a) June 27, 2006 and (b) August 18, 2006.
Figure 9. Change of estimated values in the precipitation ensemble with different number of realizations at quantile (a) 2.5%, (b) 50% and (c) 97.5%. Dimension shown on the vertical axis is 0.4 mm on (a, b); and 2 mm on (c).
Furthermore, it is realized that using the SGS method may result in poor estimation of extreme values due to inadequate field data. In such cases, the spatial variability is not fully characterized by the covariance which may lead to underestimation of the rainfall uncertainty. However, based on the case analysis in Fig.8 and Fig.9, the ensembles are able to represent the 97.5% quantile of the radar rainfall fairly good. Hence, it is believed that the method used in this research is plausible.

3.3 Impact of radar rainfall uncertainty on hydrological modeling

The rainfall uncertainty is propagated through the hydrological model using 200 rainfall realizations resulting in the same number of model simulations. Considering that radar data is available for just one year, it is decided to use stream discharge to demonstrate the impact of uncertainties in radar rainfall estimation on the simulated hydrological responses since rainfall has an almost immediate effect on stream discharge. Three discharge stations representing both small and large catchments are selected to be consistent with the previous work of He et al. [2011]. The uncertainty bound of the simulated stream discharge is shown as 95% confidence interval during the period August to December 2006. As described previously, the hydrological model runs use 2002 to 2005 as the warming up period to make sure that all simulations have the same initial conditions in the beginning of 2006. Although stream discharge responds relatively fast to the precipitation forcing, time is still needed for the cumulative effect to emerge. Hence, only results from the second half of 2006 are presented, Fig. 10.

Radar rainfall uncertainties show higher impact on the simulated stream discharge during early autumn (mid-August to September) and most of the winter. This is because the uncertainty on the discharge depends on the input rainfall amount with only subtle temporal delays. The daily variation of precipitation is also seen in Fig. 10. The dynamics of the simulated hydrograph is comparable with the observations; however, quantitatively fails to reproduce many of the small peaks. A significant number of observations fall outside the uncertainty bound indicating that the hydrological model is also affected by other factors uncertainty than those of the radar rainfall estimation. The predicted uncertainty bound only reflects the uncertainties from the estimation of radar precipitation, while the uncertainties from other model forcings, model structure and model parameters have not been included. The unsuccessful embrace of many of the observation points in Fig. 10 suggests that other uncertainty sources are critical for the simulation of stream discharge. Especially the lack of ability to simulate the small peaks is presumably due to the uncertainties in the hydrological model structure.

It is noted that the hydrological model applied in the present study has been calibrated using rain gauge station data. The radar precipitation data was not used for calibration due to the data availability of only one year. If sufficient radar precipitation data were available to carry out model calibration, the sensitivity of simulated stream discharge to rainfall input may change accordingly. This will result in slightly different hydrographs than presented in Fig. 10. However, it is believed that the model calibration will not bring significant impact on the width of the uncertainty bounds seen on Fig. 10, since it mainly affects the parameter uncertainty and has limited effect on the impact of rainfall input uncertainties.
Figure 10. Observed and simulated stream discharge with 95% confidence intervals from August to December 2006 at (a) Station no. 250021, (b) Station no. 250020 and (c) Station no. 250082. Corresponding locations of the discharge stations can be found in Fig.2. Daily mean areal precipitation is shown in each subplot based on rain gauge data.
The uncertainties from rain gauge measurement, on the other hand, are believed to have substantial influence on the uncertainties of the simulated stream flow. Taking rain gauge data as the ground truth inevitably introduces extra uncertainty to the rainfall ensemble which is considered as one of the drawbacks when using the proposed ensemble generator. The methodology in the present study takes into account the uncertainties from rain gauges in the formulation of error structure, but it is omitted when the radar pixel values are conditioned to the collocated rain gauge values. A potential solution to this problem may be to sample the measured rain gauge values based on a certain error distribution function instead of using the deterministic values. A detailed discussion of rain gauge measurement uncertainties is out of the scope of this article. However, it is expected that adding the rain gauge measurement uncertainties to rainfall ensemble used as model input will not change the baseline or the dynamics of the simulated flow but only increase the thickness of the uncertainty bound.

Comparing the hydrographs in Fig. 10, those from the upstream stations have much wider uncertainty bound than the downstream station. Therefore, radar QPE is not only more effective than traditional rainfall products for subcatchment hydrological modeling [He et al. 2011], uncertainties from radar QPE are also more evident at smaller scale. The scaling issue will be further discussed in the following.

### 3.4 Scaling issues in hydrological modeling with radar precipitation ensemble

The results suggest that the impact from uncertainties in radar rainfall estimation on the simulated stream discharge is dependent on the rainfall intensity and also the catchment scale. Since the rainfall intensity varies substantially in space and time, it is not sufficient to simply judge the degree of uncertainty by the magnitude of the uncertainty bound. Therefore, the coefficient of variation (CV) is introduced to normalize the simulated flow data, in order to compare model results from different periods of the year and at different locations. CV is defined as:

\[
CV = \frac{\sigma}{\mu}
\]

(16)

where \(\sigma\) is the standard deviation and \(\mu\) is the mean. In this case \(\sigma\) and \(\mu\) are obtained using the results from the 200 simulations of daily stream discharge. Thus, CV is a dimensionless parameter that is both a measure of the dispersion of the flow data and the prediction uncertainty.
Fig. 11 shows calculated CV at the three stream discharge stations indicated in Fig. 2 and Fig. 10 during the period August to December 2006. Three major peaks are observed at all three discharge stations: in late August, end of October and mid-December, where the highest peak occurs during early autumn. It indicates that the uncertainty propagated from radar precipitation is not only a function of rainfall intensity but that the sensitivity of the hydrological model to precipitation input uncertainty varies over season. During the dry summer season, variation (uncertainty) in precipitation will typically not result in changes (uncertainty) in runoff, because the variation in precipitation is caught by the soil moisture deficit in the root zone. In wet winter periods where the soil moisture deficit is small or non-existing, additional precipitation will generate additional runoff. In periods where the hydrological system changes from one regime to another, e.g., from dry conditions characterized by low groundwater recharge and drainage, to wet conditions small changes in precipitation may have a large impact on stream discharge. This result is in agreement with Refsgaard et al. [1983] who found that the predicted discharge uncertainty in another Danish catchment showed the highest coefficients of variation during the autumn.

It is clear that the CV of the simulated discharge decreases with increasing catchment size which confirms our previous judgment. However, it is also noticed that the strongest dynamics seen in Fig. 11 occurs at the smallest catchment area (station 250021, area: 46.5 km$^2$). The two larger subcatchments, though 10 times difference in size, appear to be quite similar and does not present as strong variations as seen at station 250021. Therefore, it is inferred that the simulated hydrological fluxes can be expected to be more sensitive to rainfall uncertainty at relatively small scale, possibly at catchment size smaller than 50 km$^2$.

![Figure 11](image-url)

**Figure 11.** Discharge uncertainty expressed as the coefficient of variation as a function of time at three discharge stations in Skjern catchment. The subcatchment areas are 46.5 km$^2$, 117.3 km$^2$ and 1054.6 km$^2$ for Station no. 250021, 250020 and 250082, respectively.
Figure 12. Delineation of Skjern catchment into 148 subcatchments. Subcatchments are delineated based on 10 m Digital Elevation Model. Subcatchment areas range from 3 to 49 km$^2$. The dominating factors of groundwater flow are marked in different colors.

To investigate the scale problem in more detail, the Skjern catchment is delineated into 148 synthetic subcatchments, shown in Fig. 12, with catchment areas that between 3 and 49 km$^2$. Mean values of the coefficient of variation for precipitation, groundwater recharge and stream flow are calculated for the period August to December 2006 at each individual subcatchment using the 200 model simulations and plotted against catchment size, seen on Fig. 13. The results show the sensitivity of simulated hydrological responses to radar rainfall estimation uncertainty at various scales. It confirms the assumption that more substantial variations are observed with decreasing catchment size. Towards the smallest scale, which is about the same size as one radar image pixel, CV of precipitation shows a range between 1.5% and 5% which indicates the magnitude of variations in the accumulated precipitation over five months. This relatively low uncertainty is mainly caused by the application of the ARNE algorithm. The uncertainty of radar QPE is quantified by comparing with rain gauge data and ARNE not only adjusts the mean bias but also reshape the spatial distribution of rainfall based on the rain gauge data points. Therefore, although the rain gauges are split into two separate groups; those used in ARNE and those used in the uncertainty analysis, much higher variation will appear if raw radar data is used without the ARNE adjustment. However, the low degree of variation seen on precipitation reflects that ARNE, as a QPE algorithm with the objective to minimize uncertainty, has satisfying performance.
When propagating the rainfall ensemble through the hydrological model, groundwater recharge, Fig. 13b, has the same or slightly higher magnitude of CV as precipitation, indicating that the degree of variation in simulated groundwater recharge mimics the trend from the input signal. The recharge is obtained by subtracting evapotranspiration from precipitation, and the simulated actual ET is very low in autumn and winter in the Skjern catchment due to the low temperature. Hence, it is not surprising to observe a close match between rainfall and groundwater recharge. However, the CV of precipitation stabilizes at 0.015 whereas the groundwater recharge levels out at 0.02. The relative difference between the two is rather significant, about 33%.

The uncertainty on the simulated stream discharge, Fig. 13c, on the other hand, is magnified from input precipitation with about three folds. Amplification of uncertainty from rainfall to stream discharge has been observed by other studies and was explained by uncertainties in the soil properties and the transpiration of the vegetation which both played an important role in amplifying the effect [Chaubey et al. 1999; Thomas and Bates 2002]. However, since uncertainty of the groundwater recharge is found to be only slightly higher than the uncertainty on precipitation this mechanism does not explain all of the amplification. The saturated zone including the drain system may also affect the uncertainty of the stream discharge.

Furthermore, the CV calculated for stream discharge is much more dispersed than that of the groundwater recharge. The runoff in the Skjern catchment is dominated by groundwater baseflow to the streams, and groundwater flow therefore plays an important role in the simulated stream discharge. The groundwater flow at each subcatchment may be generated from direct recharge and/or from boundary inflow from neighboring upstream subcatchments. Hence, a subsurface boundary inflow to recharge ratio is calculated at every subcatchment. If the ratio is greater than 1, the groundwater flow is dominated by incoming water from other subcatchments, otherwise by vertical downward recharge. It is also found that at the places where subsurface water exchange is the dominating factor, the subsurface inflow and outflow are both very large which makes the water runs quickly through the catchment boundaries with relatively little discharge generated. Therefore, the variability of simulated discharge at the through flow dominated subcatchments is very low in spite of small subcatchment sizes. The points with the calculated ratio value larger than 1 are marked in Fig.13c in light blue. If these points are neglected, the general trend of decreasing CV with increasing catchment size becomes more evident. In addition, the simulated groundwater recharge seen on Fig.13b represents a local effect reflecting the variations of the precipitation input, whereas the response from rainfall to runoff seen in Fig. 13c also includes the combined groundwater inflow from upstream subcatchments, thus no longer a local effect. This may also be the explanation to the less obvious trend in simulated stream discharge uncertainty compared to that of groundwater recharge. The subcatchments with subsurface boundary flow to recharge ratio larger than 1 are marked in Fig.12. As can be seen, these subcatchments are located in the middle of the Skjern catchment; and along the edges of catchment border, recharge always dominates.
Figure 13. Simulated hydrological responses using the MIKE SHE model with 148 subcatchments. The sensitivity of each hydrological compartment to uncertainties in radar precipitation estimation is indicated by the coefficient of variation plotted against catchment size: (a) Precipitation, (b) Groundwater recharge, and (c) Stream discharge. The dimension of the vertical axis shown on (c) is twice as much as what is shown on (a, b). The light blue marks on (c) correspond to the subcatchments indicated in Fig.12, which are the areas dominated by subsurface through flow.
It should be noted that the spatial resolution of radar QPE is 2×2 km whereas the grid size of the hydrological model is 500×500 m. Therefore, even though the radar based rainfall product represents a substantial improvement in resolution compared with the previously used rain gauge based product [He et al., 2011], the heterogeneity of rainfall at model grid scale has not been accounted for. With the development of radar technology the future radar QPE will be able to provide finer spatial resolution to comply with the requirement of hydrological models. In those cases, it is expected that the variations seen on Fig. 13 towards the smaller catchment scale will be magnified. Another factor that should be mentioned is that the geological settings of the studied catchment are mainly constituted by sandy soils. Therefore, rainfall with high intensity infiltrates almost instantly, and the stream discharge is dominated by slow groundwater discharge. In other catchments with less permeable shallow geology, e.g. clay, the impact of uncertainties of radar QPE on larger catchments might be more apparent.
4. Conclusions and summary

In the present study a statistical method to generate radar rainfall ensembles as input to a hydrological model has been developed, with the intention to assess the impact of uncertainties in radar QPE on simulated hydrological responses at catchment and subcatchment scale. Radar uncertainty realizations are obtained using coupled ordinary Kriging and sequential Gaussian simulation. This is, to our knowledge, the first time this approach is applied for quantification of radar rainfall uncertainties for integrated water resources modeling purposes. The resulting rainfall realizations are obtained by integrating the perturbed uncertainty fields with the unperturbed radar QPE fields.

Two situations are chosen to demonstrate the rainfall realizations under two different types of rainfall regimes, stratiform and convective, respectively. An empirical method is used to determine the minimum ensemble size. It is found that 200 realizations are sufficient to properly represent the error structure and also keep the computational time for hydrological simulations acceptable.

One major contribution of this study is the propagation of uncertainties of long term radar QPE data in a fully integrated and distributed hydrological model. The uncertainty bounds simulated for stream discharge show similar dynamics compared to the observed discharge but fail to include more than half of the observation points. This suggests that other uncertainty sources, e.g., from the model formulation, may play an important role to the simulated discharge.

The coefficient of variation of stream discharge is calculated using the 200 simulation results. It is found to be highest during autumn which may be explained by uncertainties in model formulation of recharge and drainage flow generation. Furthermore, it is discovered that the impact of radar rainfall uncertainties was inversely associated with the catchment size and especially significant at scales smaller than 50 km².

The scaling issues are further investigated by dividing the catchment into 148 subcatchments with sizes smaller than 50 km². CVs are calculated by using accumulated fluxes at each subcatchment. It shows that the simulated groundwater recharge has the same degree of variation compared to the rainfall input signal, whereas a three-fold-amplifying effect is found for simulated stream discharge. In addition, the scale dependency effect is found to be more evident in simulated groundwater recharge than stream discharge. This is due to the aggregating effect of runoff from the upstream subcatchments whereas the response from rainfall to recharge is rather local.

In an outlook to the future prospects, longer time series of radar data are needed to calibrate the hydrological model and to study the annual variations of the uncertainties propagated from radar rainfall estimation to the simulated hydrological results. It is also planned to generate radar QPE on a sub-daily basis, so that the temporal correlation of rainfall uncertainties can be better represented.
5. Acknowledgement

This work was part of the HOBE – Center for Hydrology (www.hobe.dk) which is funded by the Villum Foundation.

6. References


PAPER III

CALIBRATION AND VALIDATION OF A DISTRIBUTED WATER RESOURCES MODEL USING RADAR AND RAIN GAUGE BASED PRECIPITATION INPUT

Xin He, Simon Stisen, Torben O. Sonnenborg, Jens Christian Refsgaard, Flemming Vejen, Karsten H. Jensen

Submitted to Journal of Hydrology, July 2011
Calibration and validation of a distributed water resources model using radar and rain gauge based precipitation input

Xin He\textsuperscript{1,3,*}, Simon Stisen\textsuperscript{2}, Torben O. Sonnenborg\textsuperscript{2}, Jens Christian Refsgaard\textsuperscript{2}, Flemming Vejen\textsuperscript{3}, Karsten H. Jensen\textsuperscript{1}

\textsuperscript{1}Department of Geography and Geology, University of Copenhagen
\textsuperscript{2}Department of Hydrology, Geological Survey of Denmark and Greenland
\textsuperscript{3}Data and Climate Division, Danish Meteorological Institute

* Corresponding Author (xh@geo.ku.dk)

Abstract

Radar based quantitative precipitation estimation (QPE) has high spatial representation of the rainfall systems and is therefore highly suitable for distributed hydrological modeling. The present study investigates the potential of using long term radar QPE data as input to a distributed water resources model. A hydrological model developed for western Denmark is calibrated and validated using a 5-year radar QPE data series as precipitation input. Comparisons are carried out with results from a model based on rain gauge data. The model calibration suggests that the root depth of the crops is the most sensitive parameter to different rainfall products, since it controls the evaporated water fluxes and therefore the water balance. Evaluations to the optimized models show that the radar based models generate comparable results to the rain gauge based model in terms of stream discharge, groundwater head and overall water balance. However, it is insufficient only to carry out areal mean bias correction of the radar data to produce realistic rainfall estimates and that a range dependent adjustment method is required to obtain a realistic spatial distribution. The results also suggest that the advantages of using high resolution data can be hampered by the uncertainties in the QPE, especially the problem caused by accumulated errors. Furthermore, an innovative attempt is made to validate the hydrological consistencies using satellite based surface temperature data.

Key words:
Weather radar, precipitation, water resources modeling, model calibration, long term assessment
1. Introduction

Precipitation is the main driving force for hydrological processes in the terrestrial water cycle. From a water resources management perspective, accurate measurement of the temporal and spatial variation in precipitation plays an indispensable 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 measured 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; Carpenter et al. 2001; Tilford et al. 2002; Cole and Moore 2008, 2009; Gourley et al. 2010).

Radar based rainfall measurement has higher spatial and temporal resolution and 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.

The U.S. National Weather Service launched a joint venture study in 2004 aiming to explore the potential of using distributed hydrological models with radar precipitation forcing (Reed et al. 2004; Smith et al. 2004). One of the main conclusions was that the most important benefit of using the combination of a distributed model and radar data is the ability to simulate stream flows at small ungauged headwater catchments. However, it was also suggested that using a distributed modeling approach with high resolution input data may not always provide improved outlet simulations compared to lumped conceptual models with uniform rainfall input (Bandaragoda et al. 2004; Carpenter and Georgakakos 2004a, 2004b; Ivanov et al. 2004; Liang et al. 2004). One possible reason for this outcome could be that the non-linearity of distributed models magnifies the errors in high resolution rainfall data (Smith et al. 2004). Comparable results and matching conclusions have been found by He et al. (2011).

Another factor that may significantly influence the model performance is the model calibration. Intuitively, one would assume that using higher resolution data will lead to better simulation results than if lower resolution data are used. However, in reality the relation between input data resolution and model performance is not straightforward. When calibrating a distributed hydrological model, change in model parameters may result in complex redistribution of water. Very often the redistribution of water caused by model parameter change cannot be fully accounted for by using higher resolution input data (Finnerty et al. 1997). In other words the performance of the model may be more affected by inappropriate model calibration rather than deficiencies of the input data. Therefore, in many cases the accuracy of the model results depends to a large extent on the skills of the modelers rather than the type of hydrological model or input data used (Dawson and Wilby 2001; Reed et al. 2004; Ajami et al. 2007). He et al. (2011) showed that the disagreement between simulated and observed stream discharge hydrographs
can be a result of shortage in model calibration rather than deficiencies in the rainfall input data. Nevertheless, an important message learned from previous studies is that the performance of hydrological models driven by radar precipitation data can always be enhanced through dedicated model calibration (Ajami et al. 2004; Di Luzio and Arnold 2004).

The strategy for calibration of distributed hydrological models has been a topic subject to much discussion (Sonnenborg et al. 2003; Butts et al. 2004; Refsgaard and Henriksen 2004; Beven 2007; Refsgaard and Hansen 2010). It is commonly acknowledged that the calibration procedure for a distributed model is more complicated than for a lumped model where discharge at the catchment outlet commonly is the only calibration target and the number of parameters is small. Model calibration using radar based precipitation input may impose even more challenges since our knowledge of radar precipitation is far less comprehensive than that of rain gauge precipitation. Although radar precipitation offers many advantages, the quality and reliability of weather radar has always been questioned mainly because radar measures precipitation remotely and indirectly. Three major problems are of concern from a hydrological perspective: the false echoes, the mean bias and the range dependent underestimation at long distance (Kitchen and Jackson 1993; Klazura et al. 1999; Germann et al. 2006; Ciach et al. 2007). Therefore, it is recommended that radar precipitation data is adjusted by rain gauge data prior to application in hydrological modeling (Krajewski 1987; Smith and Krajewski 1991; Seo and Breidenbach 2002; Goudenhoofdt and Delobbe 2009).

Radar precipitation has been used in models subjected to calibration in several hydrological studies, but most of them are limited to short term simulations and are often event based (Sun et al. 2000; Di Luzio and Arnold 2004; Kalinga and Gan 2006). Some studies have attempted to look into the effects of seasonal variations of radar rainfall on inverse hydrological modeling, but it is very rare that both the surface water and groundwater components are taken into consideration (Guo et al. 2004; Yilmaz et al. 2005; Kalin and Hantush 2006). Calibration of a physically based, distributed water resources model using radar precipitation requires years of radar QPE data with relatively high quality. However, application of long time series of radar QPE may cause difficulties. Static noises on the radar images that are unnoticeable on small temporal scale may be accumulated in the long run and become problematic. Thus, a thorough model calibration and performance evaluation of a distributed water resources model using both radar and rain gauge based precipitation will not only benefit the scientific understanding of the physical processes in the hydrological model but also help demonstrating the potentials of radar QPE, since the hydrological model is an ideal tool to validate the accuracy and reliability of the rainfall signal.

The objectives of the present study are: (1) to prepare a long term radar QPE dataset using an advanced radar-gauge merging algorithm, (2) to carry out model calibration for a distributed water resources model using both radar and rain gauge rainfall, (3) to evaluate the model performance with respect to different hydrological compartments, and (4) to analyze the effect of spatial distribution of rainfall on simulated hydrological responses.
2. Methods

2.1 Estimation of precipitation and the QPE products

2.1.1 Radar based rainfall estimation

The present study uses a radar QPE algorithm developed by the Danish Meteorological Institute (DMI). The algorithm is based on the method developed by Michelson and Koistinen (2000) and runs on daily time step and 2 km Cartesian grid. The short name for the algorithm is ARNE and is described by He et al. (2011). In the following a brief introduction is given.

The relationship between rainfall intensity and radar reflectivity is based on the Marshall-Palmer power law (Marshall et al. 1947),

$$ Z = A \cdot R_0^b $$  \hspace{1cm} [1]

where $Z$ is the radar reflectivity and $R_0$ is the unadjusted rainfall rate. Standard parameter values are applied viz. $A = 220$ and $b = 1.6$.

The raw radar images are subjected to bias adjustment based on ground observations from rain gauges. The radar pixels with collocating rain gauges are collected as data pairs. The bias is then 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 ($R$) is obtained by multiplying an adjustment factor field to the original radar composite sum ($R_0$):

$$ R_{i,j} = R_{0(i,j)} \cdot 10^p $$  \hspace{1cm} [2]

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.

Alternatively, the more simple Mean Field Bias (MFB) correction can be applied:

$$ R = R_0 \cdot MFB $$  \hspace{1cm} [3]

where MFB is defined as the summation of all rain gauge values at a certain time step divided by the summation of the corresponding radar pixels values at the same time step. Therefore, the correction factor is unique for each radar image and will be multiplied to the entire precipitation field without local conditioning. In the present study, the MFB correction only uses rain gauge data from the considered catchment area whereas the ARNE method uses all rain gauge measurements under the radar umbrellas.
2.1.2 Rain gauge based rainfall estimation

To meet the demand for grid climate data for applications in e.g. distributed hydrological models, DMI has developed an interpolation method that projects point observation data onto 10, 20 or 40 km rectangular grids for the land area of Denmark. The interpolation is based on inverse distance weighting:

\[ Z(x_0) = \sum_{i=1}^{n} \frac{w_i(x_0)}{\sum_{i=1}^{n} w_i(x_i)} \cdot Z(x_i) \]  \hspace{1cm} \text{[4]}

where

\[ w_i(x_0) = \frac{1}{d(x_i, x_0)^p} \]  \hspace{1cm} \text{[5]}

where \( w_i \) is the weighting factor, \( x_0 \) is the grid center and \( x_i \) indicates the location of the observations; \( d \) is the distance from the center of the grid point to the observation point, and \( p \) is the power parameter which needs to be a positive number (2 in this case).

The estimated value at the center point \( Z(x_0) \) is obtained by considering the \( n \) number of observations \( Z(x_i) \) in the vicinity and their weights. Instead of using a fixed search radius, where all stations within the defined distance are included, the current 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.

The products of this method have long been regarded as the ‘standard’ climate data set in Denmark, and most of the previous hydrological modeling practices have employed these data as the model input (van Roosmalen et al. 2009; Fu et al. 2011).

2.1.3 Correction of wind induced undercatch

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. It was suggested by various studies (Larson and Peck 1974; Allerup and Madsen 1980) that the uncertainty in precipitation measurement by rain gauge is largely related to precipitation catch deficiency, which is defined as the percentage of precipitation that falls outside the opening edge of the rain gauge due to the influence of wind. The wind induced undercatch is especially significant in the winter time when solid precipitation is frequently observed.

A rain gauge catch correction model was formulated by (Allerup et al. 1997) for the Danish Hellman rain gauge. The model separates solid and liquid precipitation based on air temperature, \( T \) and wind speed, \( u \), and is given by:

\[ K = K_{\text{solid}} + K_{\text{liquid}} \]

\[ = a \cdot \exp \left[ \begin{array}{c} 0.04587 + 0.23677 \cdot u + 0.01798 \cdot T + 0.01541 \cdot u \cdot T \end{array} \right] + (1 - a) \cdot \exp \left[ \begin{array}{c} 0.00769 + 0.03433 \cdot u + 0.00101 \cdot \ln(I) + 0.01217 \cdot u \cdot \ln(I) \end{array} \right] \]  \hspace{1cm} \text{[6]}
where $T$ is air temperature at reference height (in this case 2 m above ground), $I$ is rainfall intensity, and $u$ is wind speed at the same reference height, $\alpha$ is the percentage of snow (0-100%). If the temperature is below 0 °C, it is assumed that the water phase of the precipitation is only solid, whereas if the temperature is above 2 °C only liquid precipitation is assumed. If the air temperature lies between 0 °C and 2 °C, a linear mixing model of solid and liquid precipitation is used, e.g., a temperature of 1 °C results in 50% rain and 50% snow.

### 2.2 The hydrological modeling system

In the present study hydrological modeling is based on the MIKE SHE code which is a deterministic, distributed and physically-based modeling system (Abbott et al. 1986).

The present study employs the 2009 version of the model code which includes the following model descriptions: a 2D 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 unsaturated zone describing the distribution of infiltration between evapotranspiration and groundwater recharge; a linear reservoir model for flow in subsurface drains; a 3D 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 various documentations (Yan and Smith 1994; Refsgaard and Henriksen 2004; Graham and Butts 2006).

### 2.3 Model optimization

#### 2.3.1 Optimization strategy

The overall model calibration scheme is in accordance with the guidelines given by Madsen (2003), where it is stated that the model calibration process can be divided into three key steps: First step is model parameterization where the sensitivities of the model parameters are analyzed and choices of parameters for calibration are made. Second, calibration criteria are defined and the objective functions are formulated. Third, the algorithm used for calibration is specified. The individual steps will be explained in details.

#### 2.3.2 Multi-objective calibration

Due to the complexity of the MIKE SHE model structure, multi-objective calibration is adopted which enables various hydrological processes to be optimized simultaneously (Gupta et al. 1999). Using the multi-objective calibration method situations where the model is over-conditioned to a particular water compartment due to limitations of the field data is reduced. This feature is especially beneficial in modeling where the entire hydrological cycle is considered altogether. Based on the priority of the modeling purposes, different objective functions are defined, weighted and finally aggregated,
\[ \text{Min}(\varphi_{\text{agg}}) \quad \text{and} \quad \varphi_{\text{agg}} = \sum_{i=1}^{M} w_i \cdot g_i \cdot \varphi_i \]  

where \( \varphi_i \) and \( \varphi_{\text{agg}} \) are the individual objective functions and the final aggregated objective function, respectively, \( w_i \) and \( g_i \) are the weight and the transformation factor for each objective function respectively. Transformation factors are applied in order to normalize the contributions of the objective functions regardless of the numerical values of the observations.

2.3.3 Model optimization algorithm

Model calibration is carried out using the PEST optimization tool (Doherty 2003; Doherty and Johnston 2004). PEST is a set of model independent parameter estimation program codes, which has been widely used and intensively documented in the field of hydrological research (Moore and Doherty 2005; Skahill and Doherty 2006; Christensen and Doherty 2008). The model optimization method used is the Levenberg-Marquardt Algorithm (LMA) with Jacobian matrix updating, which is a gradient based local optimization algorithm. LMA is widely used in least squares model fitting especially for solving nonlinear problems. LMA is a rather fast algorithm; however, using a gradient based optimization algorithm may increase the risk of reaching local minimum instead of a global minimum. Nevertheless, based on previous studies in the same area the initial condition given to the inverse modeling is believed to be close to the Pareto front.
3. Study area, data and model setup

3.1 General description of the Skjern River basin

The Skjern catchment is located in western Denmark next to the Ringkoebing Fjord, see Fig.1 and 2. It is one of the most heavily instrumented areas in Denmark in terms of water related research activities (Jensen and Illangasekare 2011). The catchment is approximately 3500 km². Land use consists mainly of agriculture for more than 85% while the rest is urban and forest areas.

Western Denmark is dominated by a typical maritime climate, thus the Skjern 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 a mean discharge of ca. 35 m³/s. 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 constructed to direct the excessive water to the streams.

Figure 1. Model domain for the Skjern River model. (A) shows the topography, the river network, the stream discharge stations and the rain gauge stations. (B) and (C) show the groundwater wells with hydraulic head observations before and after 2005, respectively.
3.2 Observation data

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 Fig. 2. The radars are operated by DMI and have a detection range of 240 km. Pseudo-CAPPI images of radar reflectivity at 2 km height are generated every 10 min. The maximum-pixel-value approach is adopted to generate composite scenario in order to avoid the borderline effect in the area where the radar beams intercept.

Ground based precipitation data are obtained from the rain gauge stations placed across the Skjern catchment (Fig. 1). In the present study, the 10 km rainfall grid product provided by DMI, DMI10 for short in the following, will be used as a comparison to the radar rainfall products. 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.

There are several ways to implement the rain gauge catch correction model given by Eq. [6]. The most common way is to use monthly correction factors based on 30-year statistical data collected from 1961 to 1990, the so-called ‘standard correction’ (Allerup et al., 1997). Stisen et al. (2011)
proposed a new approach where daily correction factors based on the best available observation data are calculated, named ‘dynamic correction’. It was demonstrated that the use of the dynamic correction factors largely improved the performance of the hydrological model. The main reason was that the winter temperature in Denmark has been higher in the period after 1990 than in the period used to establish the standard correction factors, and a significant decrease in snow events has been observed after 1990. As a result, precipitation is overestimated during the winter season using the standard correction factors. Therefore, it is decided to adapt the dynamic correction factors in the present study.

Observation data from stream discharge stations and groundwater monitoring wells are used (Fig. 1). The observation points are distributed quite evenly 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 rearrangement of the local authorities around the year 2005-2006, many of the groundwater data collected haven’t been reported to GEUS. As a result, most of the data collected during 2005-2006 are missing from the database. After 2007, data archiving has gradually gained momentum but still not at the level before 2005 (seen on Fig. 1.B and C).

3.3 Geology

The shallow geology in this area is dominated by Quaternary sand and gravel which forms large interconnected aquifers. Below, large Miocene sand and clay formations are found (Scharling et al. 2009). Data from groundwater abstraction wells are the main data source for building the geological model. Six conceptual 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.

3.4 Model setup

The Skjern model is based on the Danish National Water Resources Model (Henriksen et al. 2003). The simulation period is 1998-2009. Radar QPE data are available for 2005-2009, and this period is therefore used to investigate the effects of different precipitation inputs. 10 km gridded rainfall data from DMI (DMI10) are used for 1998-2004.

Model calibration is carried out for the period 2007-2009 using different rainfall forcing and the models are validated against field observations from 2005-2006. The reason for this arrangement is that a longer warm-up period is assigned to calibration (9 years) than to validation (7 years). Another reason is that the number of rain gauges in the Skjern catchment area was significantly reduced at the beginning of 2007. Thus, it is more convincing to calibrate the model using less input information, and validate it where more data are available. The model optimization process is illustrated in Fig. 3.
Figure 3. Annual mean precipitation calculated in the Skjern catchment using different precipitation estimation methods. The hydrological model optimization procedure is shown at the top of the data series.

3.5 Model parameterization

The Skjern model used for the present study has 96 parameters that can be individually specified. As a result, a rigorous parameterization becomes crucial since it is impractical to calibrate all the parameters at the same time. The general rules for model parameterization are that the selected parameters need to be sensitive to the measured hydrological variables at the observation points, they cannot be closely correlated with each other, and they should play a significant role in model conceptualization.

A subset of nine parameters were selected for optimization based on the criteria outlined above including six hydraulic conductivities, the drainage coefficient, the river-aquifer leakage coefficient, and the summer root depth of the crops, Table 1. Another 14 parameters were tied to the nine free parameters. The rest of the parameter values were fixed and values specified based on previous model applications to the catchment. Although there are several soil types in the catchment, the ratios between the root depths for the different soil types were assumed to be fixed. The same assumption applied to the ratios between the summer and winter root depth. It was found that the root depth is more important for simulating ET than the leaf area index, and it is crucial to the simulation of the water budget (Stisen et al. 2011b).
Table 1. List of parameters used for calibrating the Skjern catchment model.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter description</th>
<th>Initial value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kx_1</td>
<td>Hydraulic conductivity for Quaternary sand</td>
<td>3.94×10^{-4}</td>
<td>m/s</td>
</tr>
<tr>
<td>Kx_2</td>
<td>Hydraulic conductivity for Quaternary clay</td>
<td>3.42×10^{-7}</td>
<td>m/s</td>
</tr>
<tr>
<td>Kx_3</td>
<td>Hydraulic conductivity for Miocene quartz sand</td>
<td>8.40×10^{-4}</td>
<td>m/s</td>
</tr>
<tr>
<td>Kx_4</td>
<td>Hydraulic conductivity for Miocene mica sand</td>
<td>9.27×10^{-5}</td>
<td>m/s</td>
</tr>
<tr>
<td>Kx_5</td>
<td>Pre-Quaternary clay</td>
<td>6.73×10^{-8}</td>
<td>m/s</td>
</tr>
<tr>
<td>Kx_6</td>
<td>Top 3 m till/moraine</td>
<td>1.48×10^{-4}</td>
<td>m/s</td>
</tr>
<tr>
<td>Drain</td>
<td>Time constant for drain flow</td>
<td>2.86×10^{-8}</td>
<td>1/s</td>
</tr>
<tr>
<td>Leak</td>
<td>River-aquifer leakage coefficient</td>
<td>1.38×10^{-6}</td>
<td>m/s</td>
</tr>
<tr>
<td>RD</td>
<td>Summer root depth</td>
<td>600</td>
<td>mm</td>
</tr>
</tbody>
</table>

All free parameters are assumed to be log-normally distributed. No constrains were specified to limit the parameter searching, implying that the optimized parameter values might end up with unrealistic values. This specification were designed on purpose since this is an indication that errors may have occurred in the model input or the model formulation if unrealistic parameter values are obtained (Stisen et al. 2011b). The descriptions of the free parameters as well as the initial values are summarized in Table 1.

Table 2. List of objective functions and relative weights used for calibrating the Skjern catchment model with PEST.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Objective function</th>
<th>Unit</th>
<th>Number of observations</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2_qobs</td>
<td>Nash-Sutcliffe coefficient of daily discharge</td>
<td>--</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Fbal_total</td>
<td>Mean relative error in the total water balance</td>
<td>%</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Fbal_JJA</td>
<td>Mean relative error in the summer water balance</td>
<td>%</td>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>RMS_dyn</td>
<td>RMSE of the individual hydraulic head 2007-2009</td>
<td>m</td>
<td>236</td>
<td>1</td>
</tr>
<tr>
<td>RMS_mean</td>
<td>RMSE of the mean hydraulic head 2000-2005</td>
<td>m</td>
<td>832</td>
<td>2</td>
</tr>
<tr>
<td>ME_Ls</td>
<td>Mean error of hydraulic head in each model layer</td>
<td>m</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>
3.6 Objective functions used in the model calibration

Model calibration is carried out against observation data for hydraulic head and stream discharge. Six objective functions are defined and listed in Table 2 along with the number of observation data in each group. Note that the objective functions are equally attributed to surface water and groundwater.

In order to increase the credibility of the hydrological model, it is calibrated against time series of hydraulic head data from individual groundwater wells in 2007-2009. Mean hydraulic head data from a control period (2000-2005), where more abundant observation data are available, are also used for calibration.

As seen from Table 2, groundwater is more intensively sampled than surface water. Therefore, weighting factors need to be assigned to counterbalance the difference in the number of observation data. The present study aims to investigate the impact of different precipitation products on simulated hydrological responses. Therefore, the weights assigned to each objective function are not equal: more weight is given to the surface water than the groundwater, since the reaction from surface water to rainfall input is more rapid. Based on Eq. [7], the six objective functions are then aggregated. The initial aggregated objective function error value is 12.5, which is expected to be reduced by the end of the optimization process. The reduction of the objective function error can also be considered as a simple measure to evaluate the success of the model calibration.
4. Results and discussion

4.1 Precipitation estimated by radar and rain gauge

Two radar QPE data series each with five years of radar data (ARNE and MFB) have been prepared as input to the hydrological model. The resulting areal precipitation for the Skjern catchment is shown in Fig. 4B and 4C representing annual average values for 2005-2009. For comparison, the DMI gridded rainfall data (DMI10) is also shown in Fig. 4A. A common problem of using radar QPE in a hydrological context is that the radar estimated rainfall is biased and that the QPE has the tendency to underestimate precipitation with distance from the radar site. Furthermore, both the Roemoe and Sindal radars suffer from underestimation of rainfall due to miscalibration. This problem applies particularly to the Sindal radar. Therefore, the composite images of the Skjern catchment are dominated by the Roemoe radar using the composite strategy.

Fig. 4 shows that the systematic bias has been basically removed in the ARNE image which offers comparable rainfall intensity and spatial pattern to the DMI10 image. Considering the improvement in the spatial resolution, ARNE estimates appear to be more favorable than the DMI10 product from a hydrological modeling perspective since some of the rainfall patterns missed by the interpolation of the rain gauge data may be captured by the radar QPE. In Fig. 4A and 4B this is observed, e.g., at the high precipitation area to the north of the catchment where differences are found between the two rainfall products. The benefit outlined above may not have large influence on the simulated hydrological fluxes in medium to large scale river basins, such as the Skjern River basin, due to the spatial averaging and smoothing in the modeling processes. However, for smaller catchments with area less than 500 km$^2$ the radar QPE product is expected to make a difference (He et al. 2011).

The MFB image suggests an entirely different rainfall distribution than the other two rainfall estimates. The terraced pattern from south to north is a typical example of distance-induced underestimation. ARNE, on the other hand, successfully eliminates this type of error, thus demonstrating that the radar-rain gauge merging technique does play an important role in the construction of long term QPE products. Even though the mean bias has been removed from MFB, resulting in similar annual precipitation amount for the whole catchment area (see Fig. 3), the spatial pattern of accumulated rainfall is far from realistic. However, at the time scale of individual days or even monthly averages, this synthetic pattern is not visible and only appears on the multi-annual averages. Therefore, testing MFB as an alternative rainfall product in comparison to other rainfall products is a valid exercise but the MFB product is not recommended to operational implementation.
Figure 4. Precipitation estimates using data averaged over 2005-2009 for the Skjern catchment. (A) DMI 10 km grid product, (B) ARNE adjusted radar QPE product, and (C) MFB corrected raw radar image.

Figure 5. Calibrated parameter values using different precipitation inputs and initial parameter values. The explanations for the parameter symbols can be seen in Table 1. Scales are log-transformed for easy reading except for the root depth. DMI10, ARNE and MFB indicates calibrated parameter values using DMI10, ARNE and MFB as precipitation input, respectively, while initial parameter values from Table 1 are specified. ARNE2 and MFB2 indicate the same optimization setup as ARNE and MFB with initial parameter values from the output of PEST.1. Error bars indicate 95% confidence intervals.
When the radar QPE images are accumulated, static ground clutter is inevitably magnified. This problem seems alleviated in 2006 but very serious in 2008. These high pitches may result in 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. It is designed to create histograms using all pixels on the annual accumulated radar images. For the pixels at the top 2% of the histogram, the locations are identified and the values are recalculated using the average values from adjacent pixels. This method is used as a temporary solution to the immediate problem until a more sophisticated clutter filter is provided by either the radar manufacturer or the operator. It was shown by He et al. (2011) that this simple method removed the static noises quite effectively. However, it is not able to deal with dynamic clutters that change in space and time.

4.2 Optimized hydrological model parameters

Three model optimizations are performed using DMI10, ARNE and MFB as 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. The resulting estimates of the optimized parameters are shown in Fig. 5.

The nine calibration parameters selected can be grouped according to the hydrological features 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. Since the only difference between the three models is the precipitation input, it is not surprising to see that the largest difference occur in category (3), where especially the root depth is sensitive. The hydraulic conductivities for the subsurface sediments are very similar among different model optimizations.

Actual evapotranspiration is closely related to root depth. As seen in Fig. 3, ARNE gives higher values of estimated precipitation amount than DMI10 during 2007-2009. As a result, the calibrated root depth based on ARNE input is a factor of 2 higher than when using DMI10 in order to obtain acceptable overall water balance. Similar conclusion can be drawn for MFB. The hydraulic conductivity of the top soil \( K_{x,6} \) controls the flow rate in case the groundwater table is located at shallow depth. In the Skjern catchment more than half of the area has a groundwater table less than 3 m depth, which makes the conductivity of this geological layer sensitive to changes of precipitation. However, the optimized hydraulic conductivity of the top soil is determined by a number of factors and the trend between the optimized results is not clear.

The initial parameter values used in the model optimizations are based on past experiences. Nevertheless, the optimization may be subject to equifinality problems especially because a local search algorithm is used implying that multiple parameter combinations may lead to similar performances (Beven, 2006). In order to verify the legitimacy of the choice of initial parameter values, the parameter estimated using DMI10 were used as initial parameter values for two additional optimizations using ARNE and MFB, respectively. The results are shown in Fig. 5, ARNE2 and MBF2. In general, similar parameters values are estimated regardless of the point of departure, and the calibrated models show similar performances as well. Therefore, it is assumed that all models in Fig. 5 are valid and the simulated hydrological responses are comparable. Thus,
only the model results from DMI10, ARNE and MFB are shown in the following to avoid repetition.

The uncertainty ranges associated with the optimized parameters are different for radar and rain gauge based precipitation input. The two parameters with the largest differences in relation to rainfall products also have larger uncertainty bounds. The high uncertainty of the hydraulic conductivity of the top soil is expected since half of the model cells do not have water in this layer. Therefore, small change of water content at the dry cells will cause large uncertainty. On the contrary, it is anticipated that the high uncertainty of the root depth is caused by the uncertainty in the precipitation. Radar QPE has the risk to embrace considerable uncertainty both during measurement and post-processing. The high spatial variations of rainfall signal in radar QPE and the residual noises on the radar images are perceived by the hydrological model as larger uncertainties from the input and propagated to simulated actual ET and thus higher uncertainty in calibrated root depth. With the advances of the radar technologies, the uncertainty caused by using radar QPE can be reduced but not be eliminated.

It should be noted that model parameters can usually be better determined when additional field data are used for calibration. For instance, spatial measurements of soil moisture may help constraining the parameter optimization. Further, new remotely sensed data types from air and satellite borne sensors such as soil moisture or ET also offer promising applicability in relation to calibrating distributed hydrological models due to their distributed feature.

4.3 Model performance during calibration and validation periods

4.3.1 Stream discharge

Nash-Sutcliffe coefficients shown in Table 3 are calculated using simulated and observed stream discharge at the discharge stations shown in Fig. 1. For most stations, the model performance using DMI10 data is superior to the models using radar based QPE data during the calibration period. This is in part due to the radar data quality problem. The simulated stream discharge responds strongly to the rainfall input signals, hence an artifact in the radar image on one day can cause large mismatch compared to the observed discharge value the day after. In 2008 the Roemoe radar experienced technical problems more than half of the year and the data provided were not satisfactory. This problem affects the performance for the whole calibration period despite the advantage of high spatial resolution for the smaller subcatchments. On the other hand, the models using radar precipitation have better performances during the validation period than the calibration period mainly due to the improvement of the radar data quality. Especially in 2006 the Roemoe radar worked properly most of the year and the model performance based on ARNE in 2006 shows significant improvement. Fig. 6 shows that the model performance for simulated stream flow using ARNE data is superior for nearly all stations. This demonstrates that with a carefully calibrated hydrological model and data that wasn’t extremely biased, the QPE is able to provide good predictions outside the calibration period. This underlines the importance of minimizing errors in radar operation including a more refined clutter removal program.
Figure 6. Nash-Sutcliffe coefficient calculated for simulated stream discharge in 2006 and arranged according to catchment sizes. The simulated stream discharge is obtained from the hydrological models calibrated individually using DMI10, ARNE and MFB as precipitation inputs.

Table 3. Model performance for simulated stream discharge using different precipitation inputs during both calibration and validated period, expressed by the Nash-Sutcliffe coefficient (dimensionless).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DMI10</td>
<td>ARNE</td>
</tr>
<tr>
<td>250018</td>
<td>0.13</td>
<td>0.25</td>
</tr>
<tr>
<td>250020</td>
<td>0.71</td>
<td>0.47</td>
</tr>
<tr>
<td>250021</td>
<td>0.78</td>
<td>0.69</td>
</tr>
<tr>
<td>250075</td>
<td>0.59</td>
<td>0.42</td>
</tr>
<tr>
<td>250078</td>
<td>0.49</td>
<td>0.47</td>
</tr>
<tr>
<td>250082</td>
<td>0.73</td>
<td>0.70</td>
</tr>
<tr>
<td>250086</td>
<td>0.59</td>
<td>0.66</td>
</tr>
<tr>
<td>250097</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>250727</td>
<td>0.80</td>
<td>0.66</td>
</tr>
<tr>
<td>250728</td>
<td>0.54</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Figure 7. Simulated hydrographs for discharge stations (A) 250020 and (B) 250082 at the Skjern River. The simulated discharge is obtained from the hydrological models calibrated individually using DMI10, ARNE and MFB as precipitation inputs.
In Fig. 7 simulated hydrographs at an upstream and downstream station respectively are shown. He et al. (2011) showed that the choice of precipitation product has a significant impact on the simulated stream discharge at upstream stations using an uncalibrated hydrological model. However, the present study indicates that for models that are calibrated using individual precipitation input, the differences in the simulated hydrographs are much less obvious. This again demonstrates the potential of using radar based QPE products in long term water resources modeling despite the current data quality problem. It is also noted that in the winter between 2006 and 2007, all of the models produce discharge flow that are much higher than observations. This problem is presumably caused by the model deficiency, namely the two-layer assumption. Before the large peaks arrived, there were around 10 successive days with significant rainfall. Since the unsaturated zone 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.

4.3.2 Groundwater head

One merit of using long term radar data for calibrating a water resources model is that the impact of radar QPE on the simulated groundwater behavior can be evaluated. Table 4 shows the root mean square error, RMSE, of groundwater hydraulic head calculated for the models using both radar and rain gauge based precipitation inputs. The results are shown only for the calibration period (2007-2009) due to the previously explained data availability issue in the validation period (2005-2006). The model performances based on various precipitation inputs are actually very similar with a small preference given to the MFB product. Unlike the stream discharge, the response of groundwater head to rainfall is much slower. The catchment also acts as a filter that averages out part of the noise in the radar QPE. Therefore, the simulated groundwater head is much less sensitive to the radar artifacts and clutters. Furthermore, the stream flow measured at the discharge stations aggregates over catchments with areas of minimum 50 km$^2$, whereas the groundwater head measurements are more local. As a result, the rainfall spatial details revealed by using the radar data have demonstrated some moderate advantages over the rain gauge based precipitation estimates in the simulation of groundwater heads.

The differences between the radar and rain gauge simulated groundwater heads are calculated and illustrated in Fig. 8. The groundwater head data are extracted from Layer 3 of the saturated zone of the hydrological model and averaged for 2005-2009. Layer 3, which is located approximately 10 to 20 m below the ground surface, contains the shallow local aquifer. On average, the radar driven hydrological models predict higher groundwater elevation towards the eastern part of the catchment and lower in the western part compared to the results obtained by the rain gauge driven model. Initially, it was anticipated that the selection of the rainfall product would have limited effect on the simulated groundwater head over a time span of half a decade. However, Fig. 8 indicates that the spatial distribution of precipitation applied in the hydrological model actually has significant influence on the simulated groundwater heads. At many locations in the catchment the differences between models are more than 3 meters. In addition, it appears that the model using MFB has closer match to the model using DMI10, especially in the areas near the Ringkøbing Fjord. However, it is inconclusive which rainfall product is superior in this
regard since very few independent data are available for model calibration or validation around
the fjord area.

**Table 4.** Model performance for the simulated hydraulic head using different precipitation inputs
during the calibration period, expressed by RMSE (unit: m). Results for each model layer of the
saturated zone are presented. Weighted mean scores are calculated using the number of
observations at each layer.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Number of observations</th>
<th>DMI10</th>
<th>ARNE</th>
<th>MFB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>2.18</td>
<td>1.77</td>
<td>1.60</td>
</tr>
<tr>
<td>3</td>
<td>54</td>
<td>3.52</td>
<td>3.35</td>
<td>2.89</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>1.52</td>
<td>0.99</td>
<td>1.51</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>7.38</td>
<td>9.07</td>
<td>7.37</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>5.89</td>
<td>6.05</td>
<td>5.28</td>
</tr>
<tr>
<td>7</td>
<td>37</td>
<td>2.45</td>
<td>2.69</td>
<td>2.53</td>
</tr>
<tr>
<td>8</td>
<td>60</td>
<td>2.05</td>
<td>2.39</td>
<td>1.88</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2.42</td>
<td>1.90</td>
<td>1.64</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>3.09</td>
<td>2.79</td>
<td>2.53</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Weighted mean scores 3.50 3.74 3.21

**Figure 8.** Differences in simulated averaged groundwater head (2005-2009) by various
precipitation inputs. (A) Difference between using ARNE and DMI10, (B) Difference between
using MFB and DMI10.
4.3.3 Water balance and the overall performance

Overall water balances for the three models are summarized in Table 5 for both the calibration and validation periods. The precipitation estimated by ARNE and MFB is higher than DMI10 for the calibration period. Thus, higher ET is also observed for these two models to keep the hydrological system balanced. The ET is controlled by root depth, which is one of the parameters subjected to calibration. Therefore, even though the average precipitation estimated by ARNE and MFB is almost 5% lower in the validation period compared to the calibration period, the simulated ET is still quite large since the root depth for these two models are unchanged. It is also noticed that despite the similarity in simulated stream flow at the discharge stations, the internal flow paths are quite different between the radar and rain gauge driven models. ARNE and MFB have higher contributions from base flow, whereas in the model using DMI10 as input drain flow is more significant.

It is noticed that after merging with the rain gauges, the QPE tends to yield higher rainfall than the rain-gauge-only product. There may be several explanations to this. First, the original purpose of weather radar is to locate and trace the movement of clouds, which are used for weather forecasts. Hence, the radar processing software has the tendency to keep all the information seen on the display in order not to ignore any events. However, keeping all the information is likely to cause overestimation when the same radar image is applied in a hydrological context, since not all clouds result in precipitation. Second, the composite strategy used in the present study, which is maximum-pixel-value, may also contribute to the overestimation at pixel scale. Last but not least, the non-precipitating echoes that have not been sufficiently removed can also result in synthetic rainfall.

Nevertheless, Table 5 indicates that all three precipitation products are equally good in providing sufficient accuracy for water balance modeling. It is therefore suggested that in order to simulate the overall water balance at catchment scale, the choice of precipitation product is less important as long as the bias in the precipitation is removed.

4.4 Validation of the spatial patterns using remote sensing data

The simulated hydrological response to the differences in rainfall input signals are traditionally validated through aggregated discharge volume measured at stream outlets or groundwater head elevation observed at specific locations. However, pattern comparisons for spatial similarities of the climate variables are rather rare mainly due to the lack of spatially distributed data. New research has proposed the use of remote sensing data from space satellites as an alternative information to validate hydrological models (McCabe et al. 2005; Coudert and Ottle 2007; Stisen et al. 2011a).
Table 5. Water balance for the optimized hydrological models using different precipitation inputs during both calibration and validation period (unit: mm/year).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DMI10</td>
<td>ARNE</td>
</tr>
<tr>
<td>Precipitation</td>
<td>1055</td>
<td>1078</td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>559</td>
<td>586</td>
</tr>
<tr>
<td>Stream discharge</td>
<td>441</td>
<td>422</td>
</tr>
<tr>
<td>- Base flow</td>
<td>243</td>
<td>261</td>
</tr>
<tr>
<td>- Drain flow</td>
<td>123</td>
<td>109</td>
</tr>
<tr>
<td>- Overland flow</td>
<td>75</td>
<td>52</td>
</tr>
<tr>
<td>Groundwater pumping</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>Groundwater storage</td>
<td>-23</td>
<td>-15</td>
</tr>
<tr>
<td>change</td>
<td>Error</td>
<td>2</td>
</tr>
</tbody>
</table>

Remotely sensed images of land surface temperature (LST) from the MODIS sensors on-board the Terra and Aqua satellites are used in the present study to analyze the spatial pattern of rainfall and simulated ET. The theoretical background for this analysis is that rapid changes in LST are closely linked to rapid changes in land surface characteristics, such as changes in soil moisture content due to the large difference in thermal properties between wet and dry soils (Wetzel et al. 1984). Surface soil moisture varies both due to evaporative losses and possibly precipitation gains. Therefore, MODIS LST data from all days in the years 2006-2009 are analyzed to find incidences where the patterns in LST changes are induced by precipitation. The data are screened for cloud contamination and only images where more than 80% of the model area was cloud free and had a view angle below 45° were analyzed further. Due to the high degree of cloudiness around precipitation events, only one example was found where a prolonged dry period was followed by a single precipitation event.

The rainfall event on September 8, 2006 (day 0) is shown on Fig.9 (A.1) and (A.2) based on DMI10 and ARNE, respectively. Following the rainfall event there was a succession of dry clear sky days with high quality MODIS LST images being available on day 1 and 4 after the event. The LST change between day 1 and day 4 is shown in Fig 9 (C), which indicates a pattern similar to the observed rainfall. The simulated change in ET between day 1 and day 4 is also shown in Fig 9 (B.1 and B.2) and reflects a mixture of rainfall input pattern and model structural elements on the land surface such as soil and vegetation types. This illustrates the difficulties in spatial evaluation of model input through model output.
Although the temporal change in LST remains a relatively weak proxy for rainfall patterns, it can potentially be utilized as an independent data source for discriminating between different rainfall products with regard to their spatial patterns and for testing for hydrological consistency (McCabe et al. 2008). It was however not possible to distinguish which rainfall product performed the best on the particular event presented above since the two products were very similar. Furthermore, the nature of the precipitation mechanism may also play an important role in this type of validation. For instance, stratiform rain, which is commonly seen in western Denmark, has less distinguishable patterns than convective rainfall events.
Figure 9. Validation of precipitation and simulated ET using remotely sensed land surface temperature change. (A.1) and (A.2) Rainfall on September 8, 2006 (day 0) estimated by using DMI10 and ARNE, (B.1) and (B.2) Simulated change in ET between September 9, 2006 (day 1) and September 12, 2006 (day 4) from the hydrological model using DMI10 and ARNE as rainfall inputs, and (C) Change in MODIS LST between day 1 and day 4.
5. Conclusions and summary

The aim of the present study is to explore the potentials of using radar based quantitative precipitation estimation (QPE) to calibrate a distributed water resources model. Efforts have been made to prepare long term rainfall data series as input to hydrological simulations: two radar based rainfall products (ARNE and MFB) and one rain gauge based rainfall product (DMI10) for comparison. The hydrological model is calibrated for a period of 3 years while a separate period of 2 years was used for validation. The study reveals valuable insights in the application of continuous remotely sensed rainfall data in an off-line mode hydrological modeling framework, which otherwise cannot be obtained by using event based rainfall data.

The primary conclusion is that radar data indeed can be used to generate precipitation input to a distributed hydrological model that operates over long time scales. The model simulations by using the precipitation input from ARNE yield results for surface water, groundwater, and total water balance that are comparable to the model results driven by the rain gauge based product. The simulated stream discharge is slightly better when DMI10 served as precipitation input. Errors in simulated groundwater head show very similar results among the models tested, with small preference given to MFB based model. The simulation of total water balance suggests that all models performed equally well despite some differences in the internal flow paths. As far as the model calibration issue is concerned the conclusion is in agreement with results of other studies where the hydrological model performed better after dedicated model calibrations (Finnerty et al. 1997; Ajami et al. 2004; Di Luzio and Arnold 2004).

Model calibration shows that the hydraulic conductivity of the top soil layer and the root depth are the two most sensitive parameters to the rainfall products. Especially the root depth that controls the actual evapotranspiration yields very different values for the radar and the rain gauge based models. This result is consistent with a previous study carried out within the same catchment (Stisen et al. 2011b).

It is found that it is important to carry out the radar precipitation estimation using a range dependant radar-rain gauge merging algorithm. The average precipitation based on 5 years of raw radar data suggest that the data suffer from severe range deterioration which imposes large problems if only a simple mean bias adjustment is used. MFB is able to adjust the precipitation field to obtain reasonable mean areal precipitation but with serious local errors. The ARNE algorithm developed by DMI is proved to be able to remove most of the systematic bias and the range dependent bias, and is found to be superior to the MFB method at the time when high quality radar data is available (year 2006).

The study also indicates that the quality of the radar based precipitation plays an essential role for the performance of the hydrological model. Since the radar QPE is constituted by primarily radar data and adjusted by rain gauge data, both data sets should have acceptable quality. The uncertainty on the raw radar data should be kept as low as possible, which suggests that programs for quality assurance, such as removal of residual bias and the random clutters, should be implemented. With respect to rain gauge data the precision should be high and the density of stations should be sufficient. Presently, it is not known how many stations are required to obtain a certain level of accuracy. However, He et al. (2011) showed that the uncertainty of QPE is a
function of distance from the gauging stations and it can therefore be expected that there will be a lower limit to the number of gauges that are required to obtain a sufficient quality of the QPE.

An attempt is made to validate the spatial similarity of image patterns using satellite based surface temperature data. It is believed that this approach has some future potential if a good case example can be found, but at the moment signal correlations are weak.

Over the past years, meteorologists have been dedicated to improve the spatial and temporal resolution of the radar rainfall estimation. However, the present study addresses the problem brought by the accumulated errors of radar QPE when the data is used in a water balance context. It is implied that the advantages of using high resolution data can be largely reduced by the uncertainties in the QPE. Therefore, the quality of radar QPE data needs to be improved. Considering that the study area is at the limit of the optimal detection range of the C-band radars, a local radar installed within the catchment can be very beneficial for hydrological purposes.

6. Acknowledgement

This work was part of the HOBE – Center for Hydrology (www.hobe.dk) which is funded by the Villum Foundation.

7. References


Beven, K., 2002: Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. 189-206.


PAPER IV

PRECIPITATION INPUT
IMPACT OF PRECIPITATION SPATIAL
RESOLUTION ON THE HYDROLOGICAL
RESPONSE OF AN INTEGRATED DISTRIBUTED
WATER RESOURCES MODEL

Suhua Fu, Torben O. Sonnenborg, Karsten H. Jensen, Xin He

Published in Vadose Zone Journal, January 2011
Impact of Precipitation Spatial Resolution on the Hydrological Response of an Integrated Distributed Water Resources Model

Precipitation is a key input variable to hydrological models, and the spatial variability of the input is expected to impact the hydrological response predicted by a distributed model. In this study, the effect of spatial resolution of precipitation on runoff, recharge and groundwater head was analyzed in the Alergaarde catchment in Denmark. Six different precipitation spatial resolutions were used as inputs to a physically based, distributed hydrological model, the MIKE SHE model. The results showed that the resolution of precipitation input had no apparent effect on annual water balance of the total catchment and runoff discharge hydrograph at watershed outlet. On the other hand, groundwater recharge and groundwater head were both affected. The impact of the spatial resolution of precipitation input is reduced with increasing catchment size. The effect on stream discharge is relatively low for a catchment size above 250 km$^2$, and the effect is negligible when the entire catchment area of approximately 1000 km$^2$ is considered. In the present case the highest resolution of 500 m was found to result in the best representation of the hydrological response at subcatchment scale. Stream discharge, groundwater recharge, and groundwater head were also affected by the method for correction of systematic errors in precipitation measurements. The results underscored the importance of using a spatial resolution of the precipitation input that captures the overall precipitation characteristics for the considered catchment or subcatchment. As long as the average precipitation of the considered catchment or subcatchment is accurately estimated, the spatial resolution seems less important when the integrated response in the form of stream flow is considered.

Precipitation is the most important input to hydrological models, and an accurate representation in space and time is critical for reliable predictions of the hydrological responses. However, a complicating factor is that precipitation often exhibits large spatial and temporal variations within a catchment. Particularly it is difficult to measure or infer the spatial structure from standard gauge measurements, and an improper spatial representation constitutes a significant source of uncertainty in hydrological modeling (Berne et al., 2004). Spatial variability of precipitation impacts the predictions of the hydrological processes such as seasonal flow in streams, floods, evapotranspiration, recharge, and groundwater heads. Moreover, it is of significant importance for the closure of the water balance of a watershed or a region (Vischel and Lebel, 2007).

Over the years considerable research into the effect of precipitation on runoff response has been performed (Oblé et al., 1994; Sigh, 1997; Koren et al., 1999; Bell and Moore, 2000; Berne et al., 2004; Smith et al., 2004; Segond et al., 2007). Many studies have analyzed the impact of the spatial density of rain gauges on runoff mechanisms and found that generally the quality of the model simulations deteriorates when the density of the gauge network is reduced, see, for example, Bárdossy and Das (2008). However, studies have also shown that the significance of the impact varies widely with type of precipitation, type of model being used, general hydrological conditions of the catchment of interest, size of the catchment, and time span. Bell and Moore (2000) compared the sensitivity of basin runoff to two types of rainfall events. The authors found much greater runoff variability for convective in comparison to stratiform rainfall, noting significant dampening and reduced runoff variability during stratiform events. Koren et al. (1999) showed that the response of some rainfall–runoff models to spatial precipitation variability was scale dependent and that the level of dependency varied with different formulations of rainfall–runoff generating mechanism. Moreover, infiltration-excess type models have been shown to be more sensitive than saturation-excess type.
models (Milly and Eagleson, 1988; Winchell et al., 1998; Koren et al., 1999). Studies performed in the Walnut Gulch watershed in Arizona, across a range of sizes from 4 ha to 150 km² (Faurès et al., 1995; Lopes, 1996), revealed that the spatial rainfall distribution is important for the runoff generating mechanisms at all scales, yet the importance decreased as the scale increased due to dampening effects. Effects also vary depending on antecedent conditions when storm runoff is considered. Based on results from a 10-km² catchment in the UK, Shah et al. (1996) observed that under dry conditions, higher errors in runoff prediction (14 and 8% in peak flow and volume) were obtained to a spatially averaged rainfall input compared to the case with wet conditions (6 and 3%, respectively).

Liang et al. (2004) tested the impact of spatial precipitation resolution on the quality of model calibration based on the runoff of the 1233-km² Blue river watershed. They found that for finer spatial resolution of precipitation input, a better calibration was obtained. A critical scale was 1/8 degree (~14 km). The errors started to be significant when resolutions lower than 1/8 degree were used. Vischel and Lebel (2007) found that a threshold resolution of 20 km as a characteristic spatial scale over which the performance of the model rapidly decreased. Segond et al. (2007) compared effects of different spatial precipitation resolutions on runoff response and found that higher resolutions resulted in better predictions of peak discharge and runoff volume. The results provided by Bárdossy and Das (2008) showed that using too coarse of a rain-gauge network for estimating the rainfall input can result in poor simulation results and the hydrological model needed recalibration when different raingauge networks were used. Other studies have shown, however, that fine spatial resolution was not so important for the dynamics (Winchell et al., 1998; Booij, 2002) but important for the estimation of basin-average incoming volume (Obled et al., 1994).

The above review has shown that the effect of spatial precipitation variability on runoff response is complex, as it depends not only on the degree of spatial variability but also on catchment properties such as soils, geology, and river morphology and results from one region are not directly transferable to another.

The purpose of this study is to assess the effect of spatial resolution of precipitation on streamflow and recharge for a catchment in Denmark using an integrated and distributed hydrological model. In the analysis, degradation of the gauge network is not considered, but the same number of rain gauging stations is used for defining different spatial schemes. The effect of precipitation variability and its spatial representation is analyzed for different scales of the catchment. Further, different bias correction methods for wind effects are also included as part of the analysis.

Studied Area and Hydrological Model

The Study Area

The study area, the Alergaarde catchment in the Skjern watershed, is located at the central part of the peninsula Jutland, with an area of 1055 km² (Fig. 1). The area is relatively flat, with altitudes from 10 to about 130 m above sea level. The topography slopes gently from east to west (Fig. 2). Ninety-five percent of the soil is sandy, and the rest is characterized as clayey soil types. Land use is 60.5% agriculture, 17.4% grass, 14.0% forest, 6.3% heather, and 1.8% urban area (Fig. 3).

The weather in the area is highly dependent on the wind direction because of the proximity to both the ocean and the European continent. The dominant westerly wind results in mild winters and cool summers with variable weather and often with rain and showers. Winds from the south and east are influenced by continental weather systems characterized by low temperatures in winter and high temperatures in summer (van Roosmalen et al., 2007). During winter the dominant precipitation system is extratropical storms from directions between southwest and northwest.

Fig. 1. Map of Denmark showing the location of the Alergaarde catchment.
The frontal precipitation mechanism is enhanced by orographical effects caused by the moderate increase in surface elevation from west to east. In summer convective rain events dominate the precipitation pattern, and the most intensive rainfall events are observed from June to August with maximum daily rainfall of up to 50 to 60 mm. On average (1961–1990, Frich et al., 1997) the number of days per year with precipitation in the area with intensities larger than 0.1, 1.0, and 10 mm d⁻¹ are approximately 180, 130, and 25, respectively. The number of days with snowfall averages 33 (Laursen et al., 1999), which is almost equally distributed in the period from January through March. Mean annual precipitation is 1056 mm according to rainfall data from 1990 to 1995. Maximum precipitation is observed in autumn, and the minimum is found in spring.

The soils in the area are generally highly permeable (Fig. 3), and surface runoff (Hortonian) is normally not produced. Exceptions occur for rainfall events in situations where the soil is frozen; however, this mechanism is observed relatively seldom. In wetlands and areas near streams, surface runoff may be generated by saturation excess mechanisms. However, about 90% of the precipitation infiltrates into the soil and only a small fraction results in overland flow. The rest is evaporated to the atmosphere by interception loss. The infiltrated water is either returned to the atmosphere by evapotranspiration or recharged to groundwater. Some of the recharge is captured by the drainage system composed of tile drains, ditches, and small creeks and is transferred relatively rapidly to the streams. The remaining part of the recharge enters the aquifers and discharges to the streams downstream of the infiltration areas as base flow. The main river in the catchment is the Skjern River. The river water flows from north and east to west determined by the slopes of land surface elevation. Average discharge at Alergaarde discharge station at the outlet of the catchment is 16 m³ s⁻¹ corresponding to 480 mm yr⁻¹ (1990–1995). Hence, approximately 45% of the precipitation leaves the catchment as stream flow. Evapotranspiration and groundwater discharge out of the catchment are responsible for the remaining components in the water balance.

The description of the geology in the region is mainly based on the lithological information from water supply and oil exploration boreholes combined with seismic data. The well log information was obtained from the well database Jupiter of the Geological Survey of Denmark and Greenland (GEUS). The study area is dominated by glacial outwash sand and gravel of Quaternary age, with isolated islands of Saalian sandy till. Alternating layers of marine, lacustrine, and fluvial deposits of Miocene age underlie the Quaternary deposits. The sequence is formed by layers of mica clay, silt, and sand, together with quartz sand and gravel. Thick clay layers from Paleogene underlie the Miocene deposits, and these act as an impermeable flow boundary (van Roosmalen et al., 2007).

Hydrological Model
In this study, the hydrological modeling system MIKE SHE is selected as the model code. MIKE SHE is a spatially distributed, physically based hydrologic modeling system (Abbott et al., 1986a,b; Refsgaard and Storm, 1995). It simulates all major flow processes occurring in the land phase of the hydrological cycle. It solves the process relations for six components, each representing important physical processes in individual parts of the hydrological
cycle: (i) interception and evapotranspiration, (ii) overland and channel flow, (iii) unsaturated flow, (iv) groundwater flow, (v) snowmelt, and (vi) exchange between aquifers and rivers. In this study a two-layer water balance method is used to describe flow in the root zone. Flow in the aquifer systems is based on a three-dimensional representation of the geological stratification over which the three-dimensional governing equation for groundwater flow is solved. Flow in rivers is simulated based on river geometry, slope, and the Manning roughness factor using the Muskingum–Cunge routing method implemented in the MIKE 11 model that is used to represent river flow. Further details on MIKE SHE can be found in the model user manual (DHI, 2007).

Model Set Up

The model setup is based on the National Water Resources Model (Henriksen et al., 2003) with the modifications introduced by van Roosmalen et al. (2007). The model covers a total area of 1055 km² and is simulated on a 500 by 500 m computational grid. Ground surface elevation for grid cells is calculated from digital elevation maps on the scale of 1:25,000. The geological settings are interpreted in layers of 10-m depths based on lithological data from boreholes. A total of 38 geological layers are classified in terms of the horizontal distribution in geological units and associated hydraulic parameters. In the vertical, 16 computational layers with nonuniform thickness are used. Averaging of the hydraulic parameters within the computational layers is performed when the simulation model layers are not aligned with the geological layers. The boundary conditions for the model domain are defined from surface topography and groundwater head configuration. At a given location the same conditions are specified for all computational layers over the vertical. No-flow boundary conditions are specified along the natural groundwater divide at the upstream periphery of the catchment, while a gradient of −0.0005 is specified along a 14-km segment near the outlet of the catchment.

The stream system is digitized and bank elevations assigned to specific points along the river course. Cross-sections are assessed based on measurements at specific locations in the stream system. In the model drains represent both artificial tile drains and ditches. Additionally, the drainage description captures flow through creeks and small streams not explicitly described by the river setup. The information on the drainage system in the area is limited; therefore, the model setup is simplified using drains in the entire model area and parameterized with a constant drain level of 0.5 m below soil surface. Since the topography in the study area is very flat, drain codes that specify to which stream the generated drain flow should be routed are specified. Based on satellite data, land use is classified as grain/maize, grass, forest, heather, and urban areas (Fig. 3). The root zone is defined using two soil types, sand and sandy till. The distribution of soil type is shown in Fig. 3.

The model was calibrated in two steps (van Roosmalen et al., 2007). First, a steady-state version of the model was optimized using the automatic parameter estimation procedure UCODE (Poeter and Hill, 1999). Here the hydraulic conductivities of the groundwater zone and the leakage coefficient (conductance) controlling the magnitude of base flow were estimated. Subsequently, a transient version of the model was calibrated by trial and error against observations of hydraulic head and stream discharge from station 25.05 (Fig. 2) for the period 1991 through 1995. The parameters found in the steady-state automatic optimization were transferred to the transient model (Sonnenborg et al., 2003), while the parameters controlling the dynamics of the model response, such as storage coefficients of the saturated zone and drainage coefficients, were estimated. The calibration procedure was followed by a split-sample validation test against measurements of hydraulic head and stream discharge from the subsequent validation period. The model was calibrated using climate data from a 40-km grid network provided by the Danish Meteorological Institute (Scharling, 1999b). As will be shown below, the resolution of precipitation only has a negligible effect for the discharge simulation at the downstream station 25.05. Hence, the results presented later are expected to be unaffected by the fact that the resolution of the precipitation input used for calibration was different from the resolution of the various precipitation schemes used in the current study.

Methods

Precipitation Input

To test the effect of different spatial resolutions of precipitation on the hydrological response, daily records from 21 rain gauge stations (14 stations inside and 7 stations outside the catchment, Fig. 2) for the period from 1985 to 1995 are used. The data are split into two periods. Data from 1985 to 1989 were used as a warm-up period to reach a dynamic equilibrium, and the remaining period were used for analysis of the effects of spatial variability in precipitation.

Aerodynamic effects are the most significant source of error of point measurements of precipitation in Denmark. To correct for the undercatch caused by turbulence effects, Allerup et al. (1998) have developed correction factors that vary with the month of the year. Each rain gauge location is categorized by a shelter class according to the angle between horizontal and the surrounding obstacles, see Table 1. Shelter classes A, B, and C are assumed to result in variable degrees of measurement error due to aerodynamic effects. The most frequent shelter condition at the field site is B, followed by A and C (Table 1). The correction factors corresponding to the different classes are listed in Table 2. Shelter class A has the best shelter resulting in the lowest catch correction factors, whereas shelter class C is most unprotected and therefore has the highest correction factors.

In the study, the effect of precipitation spatial resolution and correction method on hydrological response is investigated. The six precipitation inputs presented in Table 3 are tested in the model:

P1: For each rain gauge a Thiessen polygon based on 21 rain gauge stations was defined, resulting in 21 areas of various sizes
The grid-based precipitation data set (P3) is produced by the Danish Meteorological Institute using inverse distance interpolation to a 10-km grid. The precipitation amount for each rain gauge was corrected using actual gauge catch corrections.

P2: Kriging interpolation based on 21 rain gauge stations to a square 500-m grid was conducted. The precipitation amount for each rain gauge was corrected using actual gauge catch corrections.

P3: Inverse distance interpolation to a 10-km grid was conducted. The precipitation amount for each rain gauge was corrected using actual gauge catch corrections.

P4: Kriging interpolation of available rainfall stations to a square 10-km grid. The Alergaarde catchment is covered by 22 10-km grids. The precipitation amount for each rain gauge was corrected using actual gauge catch corrections.

P5: Kriging interpolation of available rainfall stations to a square 10-km grid. The Alergaarde catchment is covered by 22 10-km grids. The precipitation amount for each rain gauge was corrected using only shelter class B corrections.

P6: Daily rainfall calculated as the average of data from the 21 rainfall stations. The precipitation amount for each rain gauge was corrected using actual gauge catch corrections.

The grid-based precipitation data set (P3) is produced by the Danish Meteorological Institute using inverse distance interpolation (Scharling, 1999a). First, interpolation of uncorrected gauge data to a 5-km grid is performed where the weights used at a particular grid are calculated as the inverse of the squared distance to the precipitation station. The interpolation routine searches for stations in four sectors and selects the nearest station in each sector. Each sector is defined according to compass directions E-N, N-W, W-S, and S-E. Hence, the interpolation may not be based on the four nearest stations. Second, the 5-km grid values are averaged to obtain 10-km grid values. Subsequently, according to the recommendations of Scharling and Kern-Larsen (2002), the grid precipitation data are corrected using correction factors of shelter class B. The Kriging data sets (P2, P4, and P5) are based on ordinary kriging of daily rainfall data where the variogram is estimated on a daily basis assuming an exponential variogram model.

Table 1. Definition of shelter classes for the Danish Hellmann gauges (Frich et al., 1997).

<table>
<thead>
<tr>
<th>Shelter class</th>
<th>Angle to horizon (°)</th>
<th>Shelter description</th>
<th>No. of gauges</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>19° &lt; α ≤ 30°</td>
<td>Well</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>5° &lt; α &lt; 19°</td>
<td>Moderately</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>0° &lt; α ≤ 5°</td>
<td>Unsheltered</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Standard values (1961–1990) of precipitation correction, k, for aerodynamic effect and wetting error as a function of shelter class and month of year (Allerup et al., 1998). Y represents annual average correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 3. Precipitation products used in the analysis.

<table>
<thead>
<tr>
<th>Description of precipitation product</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 Thiessen polygons using actual gauge catch corrections</td>
</tr>
<tr>
<td>P2 Kriging to 500-m grid using actual gauge catch corrections</td>
</tr>
<tr>
<td>P3 Inverse distance interpolation to 10 km grid using only B catch corrections</td>
</tr>
<tr>
<td>P4 Kriging to 10 km grid using actual gauge catch corrections</td>
</tr>
<tr>
<td>P5 Kriging to 10 km grid using only B catch corrections</td>
</tr>
<tr>
<td>P6 Mean of station data using actual gauge catch corrections</td>
</tr>
</tbody>
</table>

Table 4. Most significant components of the mean annual water balance from 1990 to 1995. Precipitation products are defined in Table 3.†

<table>
<thead>
<tr>
<th>P</th>
<th>ET</th>
<th>Dr</th>
<th>BF</th>
<th>O&amp;P</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1055</td>
<td>462</td>
<td>145</td>
<td>413</td>
</tr>
<tr>
<td>P2</td>
<td>1052</td>
<td>464</td>
<td>144</td>
<td>409</td>
</tr>
<tr>
<td>P3</td>
<td>1069</td>
<td>467</td>
<td>151</td>
<td>412</td>
</tr>
<tr>
<td>P4</td>
<td>1066</td>
<td>465</td>
<td>151</td>
<td>413</td>
</tr>
<tr>
<td>P5</td>
<td>1096</td>
<td>467</td>
<td>162</td>
<td>424</td>
</tr>
<tr>
<td>P6</td>
<td>1095</td>
<td>462</td>
<td>148</td>
<td>409</td>
</tr>
</tbody>
</table>

† P, precipitation; ET, actual evapotranspiration; Dr, drain flow to river; BF, base flow to river; O&P, groundwater outflow and pumping.

In Table 4 the mean annual precipitation for each of the six precipitation products are listed. Note that the differences between the precipitation sums are small.

**Evaluation of Hydrological Response**

Daily stream-flow data from three river discharge stations (25.05, 25.08, and 25.24; Fig. 2) for the period 1990 through 1995 are used for comparing the responses of the hydrological model to different precipitation schemes. In addition, river discharge from 13 points representing different catchment areas is extracted from the MIKE 11 model. Groundwater head elevation data from two wells in the catchment (Fig. 2) are also used in the analysis. The daily head elevation data of the two wells are extracted from the simulation results of MIKE SHE model. Further, the mean daily head elevation at a shallow (layer 2) and a deep (layer 14) layer are also extracted for the whole watershed to show the effect of precipitation input on groundwater elevation.

**Performance Criteria**

The Nash–Sutcliffe model efficiency coefficient ($E_1$) (Nash and Sutcliffe, 1970) is used as an objective performance criterion when evaluating the ability of the model to simulate streamflow at multiple sites. The model efficiency index is computed as
\[ E_1 = 1 - \frac{\sum_{t=1}^{T} (Q_{\text{obs}}^t - Q_{\text{sim}}^t)^2}{\sum_{t=1}^{T} (Q_{\text{obs}}^t - Q_{\text{obs}}^t)^2} \]  

where \( Q_{\text{obs}}^t \) and \( Q_{\text{sim}}^t \) are the observed and simulated river discharges at time \( t \), respectively, and \( Q_{\text{obs}} \) is the average observed discharge.

To compare the effect of different spatial precipitation resolutions on river discharges, the simulated river discharge for P1 is considered as a reference, and a modified Nash–Sutcliffe coefficient (\( E_2 \)) is defined for the comparison analysis

\[ E_2 = 1 - \frac{\sum_{t=1}^{T} (Q_{\text{P1}}^t - Q_{\text{sim}}^t)^2}{\sum_{t=1}^{T} (Q_{\text{P1}}^t - Q_{\text{P1}}^t)^2} \]

where \( Q_{\text{P1}}^t \) is simulated river discharges for precipitation input P1 and \( Q_{\text{sim}}^t \) represents simulated discharges for one of the alternative precipitation models, respectively. \( Q_{\text{P1}} \) is the average discharge for simulation P1.

**Results**

**Precipitation Spatial Variability**

Based on station data corrected using site specific correction factors, the coefficient of variation of daily rainfall has been computed (Fig. 4). The CV varies from 0.05 to 4.59 with an exponential decrease for larger daily rainfall amounts. If only rainfall events larger than 1 mm d\(^{-1}\) are considered, 62% has a CV less than 0.5, and if only rainfall events above 10 mm d\(^{-1}\) are considered 95% has a CV less than 0.5. This shows that high-intensity precipitation events apply more uniformly over the catchment, whereas low-intensity events are associated with higher spatial variability.

The spatial distribution of average annual precipitation based on measurements from the 21 stations for the period 1990 to 1995 is depicted in Fig. 5. The Thiessen polygon method, P1, using site specific gauge catch correction factors yields mean annual precipitation between 900 and 1100 mm, with a relatively homogeneous distribution in the southern part of the study area, whereas more spatial variability is found in the northern and eastern parts. Minimum annual rainfall is found in the northeast corner of the catchment, while maximum is observed in the center and in the northwest part. The ratio of maximum to minimum mean annual precipitation is 1.23, and the CV of annual values is 0.05.

Figure 5 also depicts the spatial configuration of the other precipitation schemes as represented by the annual means for the period 1990 through 1995. The largest variation in precipitation is found for P1 with up to 200 mm in difference between subregions of highest and lowest rainfall. It is interesting to see that the different
resolutions used for deriving the data sets P2, P4, and P6 result in considerable differences in spatial patterns. P2, with a grid size of 500 m, has a higher spatial variation of precipitation with differences up to 160 mm, while the larger grid size used for P4 results in smaller spatial variation in precipitation, with a difference in precipitation of only 80 mm.

The spatial difference in precipitation found for P3 and P5 is the effect of different interpolation methods. The kriging interpolation, P5, yields higher precipitation than P3 using inverse distance interpolation. This is partly explained by the truncation performed in the kriging method where negative precipitation estimates are set equal to zero. As a result the average precipitation found using the kriging method is higher, which is also seen from Table 4. However, the spatial distribution in precipitation of P3 and P5 is similar. The resulting spatial variation of P3 and P5 is notably smaller (<100 mm) than P1, even though the three products have similar resolution. Reasons for this difference are that (i) correction of grid precipitation for P3 and P5 is performed using only shelter class B and the variation in correction factors is therefore not considered and (ii) grid values are found from interpolation (both kriging and inverse distance) averaging out spatial heterogeneity.

**Effects on the Integrated Catchment Response**

Table 4 shows the simulated annual water balance for the whole catchment. Evapotranspiration and base flow to rivers vary only marginally among the six simulation scenarios, while the drain to river components of P3, P4 and P5 are slightly higher than those of P1, P2, and P6 because of higher precipitation. However, the effect of precipitation input on the overall hydrological responses and on the integrated water balance is relatively small.

The simulated discharge hydrographs at station 25.05 at the outlet of the catchment (Fig. 6) show no apparent differences for the six different precipitation resolutions. Recall that we adopted a calibration from another study and we made no attempt to improve simulations of, for example, low-flow situations.

The simulated peak flows of P5 are slightly higher than the others while for lower flows the simulated discharges are almost the same. These results are also reflected by the Nash–Sutcliffe coefficients ($E_1$) for the six scenarios P1, P2, P3, P4, P5, and P6 which are found to 0.27, 0.25, −0.24, −0.21, −0.78, and −0.60, respectively. The largest impact is seen in the relatively wet period from December to May, where the largest differences between correction factors of the different shelter conditions are found (Table 2).

The difference in stream discharge between P2, P4, and P6 shows the effect of spatial precipitation resolution. The discharge of P2, with a spatial resolution of 500 m yields a result that is very different from P4 and P6. For the current discharge station the average precipitation increases with grid size, resulting in higher peak flow rate. The effect of resolution is also recognized in the Nash–Sutcliffe coefficients of 0.25, −0.21, and −0.60, respectively, for P2, P4, and P6.

When comparing the results of P4 to P1, both describing the precipitation at approximately the same resolution, it is seen that resolution is not the only factor that affects the results. The interpolation and averaging performed in the 10-km kriging product smears out local precipitation heterogeneity patterns and impacts the discharge simulations by the hydrological model.

The $E_2$ values for station 25.24 presented in Fig. 8 shows that resolution (compare P2, P4, and P6), the procedure for correction of measured precipitation (compare P4 to P5), and the interpolation method (compare P3 to P5) are important for the simulation of discharge response. The three precipitation stations located in northeastern part of the catchment upstream of station 25.24 are classified as shelter condition A, corresponding to relatively small correction factors (Fig. 2). When deriving the grid precipitation schemes P3 and P5, shelter class B is a standard routine used for all stations without consideration of the local shelter classes. The shelter class issue also explains why the largest differences between precipitation models are observed in the winter period since the correction factors of the different shelter classes vary mostly in

![Fig. 6. Stream discharge at station 25.05 at the outlet of the catchment.](image-url)
this period. Based on the simulation results of stream discharge for station 25.24, it seems that the true precipitation is described most accurately by methods using relatively high resolution and site-specific catch correction factors.

For station 25.08 with a watershed area of 82 km² (Fig. 7b), similar results are obtained with relatively large and small effects for the winter and summer seasons, respectively. As for discharge station 25.08, P5 produces the highest discharges due to catch correction and interpolation methods. Comparison of the discharges using P2, P4, and P6 shows that apparently the resolution does not affect the results at this station, and the same are found from the $E_2$ values of the three results (Fig. 8).

At discharge station Q1 (Fig. 2), representing a catchment area of 431 km², the effect of precipitation model is relatively low (Fig. 7c). P5 deviates slightly from the other precipitation models; however, the impact of precipitation input is not significant. The effect is almost eliminated at station 25.05, with a catchment area of 1055 km², as shown in Fig. 6 and 7d. This station effectively integrates rainfall heterogeneity within the catchment, and no effect is manifested of either resolution or correction method.

The interrelation between precipitation resolution and catchment size is also demonstrated in Fig. 9, which plots the changes in the Nash–Sutcliffe coefficients ($E_2$) for 13 discharge points randomly distributed in the river system. $E_2$ was calculated using the results from P1 as reference. $E_2$ is independent of precipitation resolution for catchment areas larger than 250 km², while for sizes below this value the $E_2$ coefficient shows sensitivity to the spatial resolution and the correction–interpolation method. The discharge stations with most variability in $E_2$ are all located in the northern area characterized by relatively large differences between the precipitation models.

**Effects on Groundwater**

In Fig. 10 the spatial distribution of recharge for the various precipitation schemes is illustrated. Since virtually no overland flow is present in the catchment, spatially averaged annual groundwater recharge corresponds to the sum of drain flow, base flow, and the export of groundwater out of the catchment presented in Table 4. From the simulation based on P6, representing spatially averaged precipitation, it is evident that recharge is affected by parameters other than precipitation. Recharge varies due to variations in soil type and land use, with the latter factor having the highest impact. Areas with relatively low recharge are found where the land cover is forest. The dominating forest type in the area is conifer, which is assumed to have higher evapotranspiration than the other land use types, mainly because of higher interception loss, and thus a resulting lower recharge. However, the P1 simulation in particular also demonstrates that variability in precipitation is reflected in the recharge as the polygon shapes are clearly distinguishable.
on the recharge patterns. The annual recharge in this scheme varies from less than 450 mm to approximately 800 mm.

The relative effect of precipitation product and associated recharge on the shallow mean daily groundwater heads is shown in Fig. 11. The results are illustrated as changes in mean daily heads for scenario P1, P2, P3, P4, and P5 relative to the simulation results based on P6. Spatial heterogeneity in precipitation is seen to have an impact on groundwater head distribution, and as expected, the spatial pattern of changes in hydraulic heads is consistent with the spatial pattern in precipitation. However, it is interesting to see that the impact on heads spreads outside the area of influence from precipitation. For example, the Thiessen polygon product, P1, with lower precipitation in the eastern part of the catchment (Fig. 5) affects the groundwater heads in a much larger area surrounding the polygon. The effects are most pronounced for P1 and P2, where the differences compared to P6 ranges from −1.5 and up to 0.6 m.

The effect on the deep groundwater (not shown) is comparable to the results for shallow groundwater, yet the distribution of the differences is smoother. The results show that the precipitation model is important for both shallow and deep groundwater.

Figure 12 shows the effect of precipitation model on the temporal response on groundwater head at location 87.57 (Fig. 2) representing shallow and deep groundwater, respectively. The differences between the six models are generally within 2 m, with a tendency for increasing differences during winter. The impact of recharge events is clearly seen for all the models, and the responses of the different models are similar. In Fig. 12b the response of the deep groundwater at the same location is shown. Here there are differences of up to 2 m between the P1 and the P6 models due to significantly different precipitation inputs. Contrary to the shallow groundwater, the effect of the individual recharge events is not seen at depth, resulting in a much smoother and delayed response over the season. In Fig. 13 the model response from well 105.374 is shown together with monthly measurements from shallow groundwater. At this location the difference between the models is relatively small. Only P5 produces a significantly higher groundwater level because of the relatively high precipitation and recharge in this area. All models underestimate the measured hydraulic heads; however, the bias is within a few meters, which is considered satisfactory. Although the groundwater dynamics are not captured well by the monthly data, they indicate that the fluctuations of models are slightly overestimated.

**Discussion and Conclusions**

In this study the effects of resolution and measurement error correction method of precipitation have been analyzed. The results presented show that precipitation input may have major impact on the response of a distributed hydrological model. Stream discharge, groundwater recharge, and groundwater heads may be affected by both the resolution of the precipitation input and the method used for measurement error correction. For stream discharge Nash–Sutcliffe coefficients between 0.27 and −0.78 are found for the different models for station 25.24, and for groundwater heads differences ranging from −1.5 to 0.6 m are found. The results underline the importance of using a spatial resolution of precipitation that is representative for the scale of interest. The appropriate spatial resolution will depend on the characteristic scale of the precipitation structure in relation to the problem scale.

In the present case the impact of resolution of precipitation input is reduced with increasing catchment size. This result is
supported by studies of Obled et al. (1994), Winchell et al. (1998), and Booij (2002). The effect on stream discharge is relatively low for catchment sizes above 250 km², and the effect is negligible for catchments larger than 1000 km², which indicates that high resolution of precipitation input cannot improve the simulation results of streamflow for large-scale models. Generally, the catchment area required for averaging out the effects of precipitation variability and differences in shelter classes depends on the degree of heterogeneity of the precipitation distribution. The area studied is characterized by a relatively homogeneous distribution in precipitation. Ground surface elevation is relatively flat, with
moderate increases towards the east, resulting in small orographical effects. The precipitation mechanism in the area is dominated by frontal rain that results in relatively homogeneously precipitation distribution (Fig. 4). However, in areas characterized by larger heterogeneity in precipitation distribution, it is expected that the area required to integrate out the impact may be larger.

The study also shows that it is important to use the local shelter conditions of the precipitation stations when correcting measurement errors. From a practical point of view it is attractive to generate grid precipitation without consideration of the local shelter conditions. The produced grid product can subsequently be corrected using the correction methods preferred by the user, which makes it robust both for different applications and for changes in correction factors. The correction factors are determined based on experimental data that are hard to generate and associated with substantial uncertainty. During the last 30 years the correction factors used in Denmark have been modified twice (Allerup and Madsen, 1979; Allerup et al., 1998), illustrating the difficulties of assessing such factors. Additionally, the shelter conditions at a precipitation station are likely to change with time as vegetation and other obstacles surrounding the station may change. Hence, the correction factors may change with time, and this adds to the problem of estimating grid precipitation without consideration of the shelter conditions at the individual station used in the interpolation scheme.

The study also shows that spatial heterogeneity in precipitation impacts the local recharge and thus the local groundwater head dynamics. While the impact of precipitation variability on streamflow is dampened due to the smoothing effect of the catchment, this effect is less dominant for recharge and groundwater heads.

The effects of precipitation input are so dominant that it could potentially impact the estimates of model parameters when the hydrological model is calibrated. To compensate for errors using a low resolution precipitation input or a precipitation input with inaccurate corrections for measurement errors, the model parameters controlling, for example, evapotranspiration or groundwater flow will be biased. This results in a hydrological model that is likely to lose predictive capability, especially in situations with significant changes in meteorological input, such as when used for climate change impact studies.

On the basis of this study, we recommend using station specific catch correction of measured precipitation. Further, it is important to avoid spatial heterogeneity being averaged out by inappropriate use of interpolation schemes. If the Thiessen polygon method is used, averaging effects are avoided, but at the expense of unrealistic jumps in rainfall when crossing the boundaries between the polygons. These unphysical structures are propagated through the hydrological model to groundwater recharge and to some extend also to groundwater levels. Therefore, interpolation is necessary, but to a scale that hinders too much averaging of the spatial heterogeneity. In the present case kriging interpolation to a 500-m grid has allowed us to meet these requirements.

References


Allerup, P., and H. Madsen. 1979. Accuracy of point precipitation measurements, Klimatologiske Meddelelser S. Danish Meteorological Institute, Copenhagen.


