Influence of Soil Heterogeneity and Spatial Discretization on Catchment Water Balance Modeling

This study investigated the impacts of the spatial variability of soil hydraulic properties and the effects of spatial discretization on the water balance in a fully coupled system. The integrated surface–subsurface, three-dimensional, finite element model HydroGeoSphere was applied to the forested Wüstebach basin (27 ha) to simulate water fluxes. The fully coupled flow simulation model was applied to the headwater catchment at two different spatial resolutions (25 and 100 m). The change in spatial resolution required an aggregation of the soil map, which influenced the water fluxes and the spatial patterns of soil moisture. The nonlinear relationship between soil moisture and transpiration caused the spatial aggregation of soil moisture to have a larger effect on the water balance than did aggregating the soil hydraulic properties. In addition to the total discharge, the effects on the spatial patterns of the simulated soil moisture were also investigated. The results show that aggregating soil hydraulic properties results in lower uncertainties than does using a coarser discretization. This can be explained by the nonlinearity of the relationship between soil moisture and evapotranspiration.

In the past, hydrologic modeling has focused on estimating water balance components integrated across a catchment, whereas nowadays, spatially distributed descriptions of water fluxes and state variables are in demand because of the growing interest in environmental change. The management of water quantity and quality (e.g., flood forecasting, water pollution, and sediment transport), which is influenced by land-use change and anthropogenic activities, is gaining increasing importance. Hydrologic models are well-established and widespread tools used to analyze these water-related problems (Loucks et al., 2005).

Hydrologic models must be flexible because the scale of application may differ both temporally and spatially. Scaling is an important issue in hydrologic modeling because different scales must be considered. For example, if one is interested in water fluxes within a catchment, then sampling can be performed at the laboratory scale (soil samples of a few hundred cubic centimeters); however, the processes may take place on a much finer (pore scale, millimeters) or larger (square meters) scale. A model may be applied at the field scale (hectares) or the catchment scale (square kilometers), and its evaluation is often performed using discharge measurements, which only provide integral information concerning the investigated catchment. The validity of the hydrologic variables and the physical laws is related to the operative scale of the hydrologic models (Dehotin and Braud, 2008). The distribution of the properties, for example, the soil hydraulic properties, soil moisture conditions, and their stochastic nature, are strictly related to the scale of observation (Lacey and Grayson, 1998). Models developed for a given scale are not necessarily valid at other scales (Blöschl and Sivapalan, 1995; Blöschl et al., 1997).

Models developed for microscale applications require detailed data, and even if meso- and macroscale models are based on the same physical concepts, the required data are rarely available (Grayson et al., 2002). A more flexible approach for large-scale modeling is required because, at this scale, the parameters are generally estimated and not measured and it is questionable whether models based on microscale concepts can be applied at this scale (Flügel, 1995; Bongartz, 2003). Moreover, growing ecological interest requires greater attention to the quantification of all soil–water–plant processes and their interactions (Green et al., 2004).

Water flow in the vadose zone, solute transport processes, and energy fluxes are modeled with soil hydraulic functions that are usually obtained from point measurements.
(Hopmans and Schoups, 2005). Measured samples are on the order of a few centimeters, and the distance between samples is typically more than several meters (Hopmans et al., 2002). Large-scale information is obtained from small-scale measurements via upscaling. A review of the available approaches in the literature concerning the upscaling of soil hydrologic processes and hydraulic properties was presented by Vereecken et al. (2007). In that review, upscaling was defined as the “processes that replace a heterogeneous domain with a homogeneous one in such manner that both domains produce the same response under some upcaled boundary conditions.”

Nowadays, the application of integrated surface–subsurface models to analyze real catchments is common (Partington et al., 2009). Those models solve the Richards equation in a defined set of nodes, and increasing the number of nodes causes an increase in the required computational power; therefore, the accuracy from fine model discretization and the required computer power must be balanced. The nonlinearity of the processes involved does not allow a general relationship between accuracy and model discretization to be obtained. Instead, these relationships must be determined separately for each problem to be solved. Apart from the typical problems of temporal and spatial discretization, the effects of numerical instabilities must be considered (Vogel and Ippisch, 2008). A recent study showed that temporal discretization of the input parameters (e.g., temperature and precipitation) does not have a large impact on the water balance (Kling and Nachtnebel, 2009); however, accurate and stable solutions do require a small time step (Lal, 1998). Time steps should be chosen to optimize the solution in regard to the stability and the computation time, and an adaptive time step has been recommended as a robust technique for ensuring model efficiency (Panday and Huyakorn, 2004). The temporal and spatial discretization should be optimized not only in terms of the water balance error but also in terms of numerical instability. Fine spatial discretization yields improved simulation results. Zhao et al. (2009) showed that, for grid cell sizes of 50 m up to 1 km, there is a high variability of the yearly water balance. In reference to the accuracy of the overland flow, Jaber and Mohtar (2003) showed that the upscaling slope and Manning’s coefficient have a larger effect on the water balance than does aggregation of the rainfall.

Normally, the horizontal discretization is on the order of meters up to hundreds of meters and the vertical discretization is on the order of centimeters. In the unsaturated zone, the vertical discretization near the soil surface must be fine, on the order of centimeters, but it can increase to up to 1 m in the saturated zone (Vogel and Ippisch, 2008; Downer and Ogden, 2004). Refining the vertical and horizontal discretization yields more stable and realistic solutions but requires increased computational time. Coarsening the resolution may decrease the required computer power but has undesirable impacts on the investigated water fluxes and state variables.

The importance of calibrating and validating the spatial distribution of state variables, in addition to the more traditional discharge data, has been widely highlighted (Grayson and Blöschl, 2000; Grayson et al., 2002) but is often limited because of lack of measurements. If spatial information is available, then quantitative methods for comparison should be used (Wealands et al., 2005).

The goal of this study was to evaluate the influence of spatial discretization on model results and compare it with the effect of upscaling of the soil parameters considering all components of the water balance and the spatial patterns of soil moisture. To achieve this goal, a fully integrated surface–subsurface flow model, the three-dimensional numerical code HydroGeoSphere (Therrien et al., 2007), was applied to the 27-ha Wüstebach basin, which is a tributary to the Erkensruhr River in Germany.

The model was applied at two different spatial discretizations to evaluate its ability to simulate the water fluxes at different scales; the basin used for the model was discretized with 100- and 25-m resolutions, which also affected the required computational times. The effect of upscaling the soil hydraulic properties and the effect of a coarser spatial discretization were differentiated. In contrast to other studies, this investigation concentrated on a small headwater catchment and focused on the discretization effect on the actual evapotranspiration. The impacts on the spatial patterns of soil moisture were quantified using a cell-by-cell comparison method that was based on a contingency table and is expressed in terms of kappa statistics.

| Fully Coupled Flow Simulation Model Generality |

Models for water and solute transport at the surface and the subsurface are helpful tools for understanding physical processes and for validating scientific hypotheses. While analytic models are only available for limited situations, numerical models exhibit broader applicability to real situations. Surface and subsurface water flow models can be differentiated into three classes: the uncoupled, the iteratively coupled, and the fully coupled models (Ebel et al., 2009; Furman, 2008).

A variety of numerical codes have been developed to simulate coupled surface–subsurface flows under variably saturated conditions, such as InHM (VanderKwaak, 1999), ParFlow (Kollet and Maxwell, 2006), and HydroGeoSphere (Therrien et al., 2007).

These models take into account all components of the hydrologic cycle, and the governing equations for the surface and the subsurface domains are simultaneously solved in a single combined matrix (Rooij, 2008). This is in contrast to the previous generation of models, in which the equations were solved separately for each domain and were eventually assembled via iteration (Holvoet, 2006).
A detailed overview of the available models was provided by Rassam and Werner (2008), who reported the main characteristics of each model.

### Governing Processes

A rigorous description of the physical laws is required to simulate the flow and solute transport processes in three dimensions in both the saturated and the unsaturated zones while considering the general boundary conditions (Panday and Huyakorn, 2004). The model applied in this study was the three-dimensional, control-volume, finite element model HydroGeoSphere, which was developed by the University of Waterloo (Waterloo, ON, Canada), Laval University (Québec City, QB, Canada), and HydroGeoLogic Inc. (Herndon, VA). Flow and mass transport in the porous medium are simulated in three dimensions using the mixed form of Richards’ equation for variably saturated subsurface flow, while the overland flow is simulated using the two-dimensional diffusion wave approximation of the Saint Venant equation (Panday and Huyakorn, 2004). The HydroGeoSphere model has been successfully applied at a large range of scales, from the local scale (Partington et al., 2009; Gleeson and Novakowski, 2009; Goderniaux et al., 2009) to the continental scale (Lemicieux et al., 2008a,b). A detailed description of the physical processes that are simulated in the HydroGeoSphere model can be found in Panday and Huyakorn (2004), Li et al. (2008), and in the user manual (Therrien et al., 2007).

Surface flow is an important component of the hydrologic cycle, which governs flow to and from the subsurface, channel networks, rivers, lakes, and reservoirs. For the purpose of our work, it should be noted that the soil retention curves are given by the van Genuchten function (van Genuchten, 1980), and for the surface flow, the roughness can be expressed in terms of Manning’s coefficient.

For the surface and subsurface flow, first-type (Dirichlet) or second-type (Neumann) boundaries conditions can be specified, which can vary in space and time (HydroGeoLogic, 2003). Evapotranspiration is an important component of the water balance, and for this particular reason, we provide its description below.

### Treatment of Evapotranspiration

Canopy interception and evaporation, plant transpiration, and soil evaporation are modeled in both the surface and subsurface flow domains. Actual evapotranspiration depends on the potential evapotranspiration, $E_p$, and on water availability. The value of $E_p$ may be derived from measurements or computed from vegetation and climatic factors (solar radiation, wind, air humidity, and temperature) using the Penman–Monteith equation (Monteith, 1981).

The computation of actual evapotranspiration is performed in two steps. In the first step, the canopy interception ($E_{can}$) is evaporated at the potential rate. When the interception storage is empty, water is removed in the second step from the root zone along with transpiration. Evaporation is related to the area that is not covered by plants, which is estimated from the leaf area index. Evaporation occurs from defined upper soil layers (Li et al., 2008).

Water uptake for transpiration occurs within the root zone, which may be above or below the water table. The rate of actual transpiration ($T_p$) was estimated using the following equation, which distributes the net capacity for transpiration among various factors (Kristensen and Jensen, 1975):

$$T_p = f_1 (LAI) f_2 (\theta) RDF (E_p - E_{can})$$  \[1\]

where $f_1(LAI)$ is a linear function of the leaf area index, $f_2(\theta)$ is a function of nodal water content (cf. Eq. [2] below), RDF is the time-varying root distribution function, and $E_{can}$ is the canopy evapotranspiration. The function $f_2$ describes the correlation of $T_p$ with the moisture status in the root zone and is a simplification of the function that accounts for root processes (Kristensen and Jensen, 1975).

It is worth explaining in more detail the specific role that the function $f_2(\theta)$ has on transpiration. The function $f_2(\theta)$ has a large impact on the actual $T_p$ because of its multiplicative function in Eq. [1]. The moisture content dependence is expressed as

$$f_2 (\theta) = \begin{cases} 0 & \text{for } 0 \leq \theta \leq \theta_{wp} \\ 1 - \left( \frac{\theta_{fc} - \theta}{\theta_{fc} - \theta_{wp}} \right) C_3/E_p & \text{for } \theta_{wp} < \theta < \theta_{fc} \\ 1 & \text{for } \theta_{fc} < \theta < \theta_{o} \\ \frac{\theta_{an} - \theta}{\theta_{an} - \theta_{o}} C_3/E_p & \text{for } \theta_{o} < \theta < \theta_{wp} \\ 0 & \text{for } \theta_{an} \leq \theta \end{cases}$$  \[2\]

where $C_3/E_p$ is a dimensionless ratio of fitting parameters, with $C_3$ being a function of the soil type and the root density, and its increase will reduce the influence of the low soil moisture content and vice versa; and $\theta_{fc}$, $\theta_{wp}$, $\theta_{o}$, and $\theta_{an}$ are the moisture contents at field capacity, the wilting point, the oxic limit, and the anoxic limit, respectively. They are computed as $\theta_{(subscript)} = \phi S_{(subscript)}$, where $S$ is the relative saturation and $\phi$ is the porosity of the soil. The relative saturation is the ratio of the pore volume that is filled with water and therefore varies between zero for dry and one for saturated soil.

The transpiration is zero when the water content is below the wilting point. Between the wilting point and field capacity, a linear dependency $(C_3/E_p = 1)$ between soil moisture and the water stress factor $f_2$ is assumed. Between field capacity and the oxic moisture content, no water stress occurs, whereas the water stress increases from oxic to anoxic water contents because the roots are inactive due to the lack of aeration (Feddes et al., 1978).
evapotranspiration is estimated for the elements of each layer and is assigned to the nodes based on the contribution of each element to the nodes. For each time step, the total transpiration is calculated as the sum of the transpiration from each node in the root zone. In this model, the water stress in one layer is not compensated for by increasing water uptake in one of the other layers.

### Grid Size Effect on Evapotranspiration

The nonlinear function $f_2(\theta)$ (Eq. [2]) plays an important role in the upscaling process; therefore, its behavior under an artificial condition was analyzed. To explore an example of what can happen in the upscaling process, this function has been characterized for a single group of elements and wet conditions ($\theta > \theta_f$). For the following example, the oxic and anoxic limits were fixed at $0.85$ and $0.95$ of relative saturation, respectively, and $C_3/E_p = 1$. In Fig. 1, a generic element $A$ and a possible mesh refinement are shown. If $a_i$ is the area of the $i$th generic fine element, the area $A$ of the coarse element $A$ can be written as $A = \sum_{i=1}^{n} a_i$, with $n$ equal to five, according to the example.

We assumed a homogeneous porosity, that each element had an arbitrary relative saturation, that $\theta_1 = \theta_3 = \theta_4 = 0.85$, and that $\theta_5 = 1.0$, such that the relative weight of the function $f_2(\theta)$ can be defined as

$$f_2(\theta) = \sum_{i=1}^{n} a_i f_2(\theta_i)$$  \[3\]

where $f_2(\theta_i)$ is, from Eq. [2], equal to 1 for $i = 1, 2, 3$ and equal to 0 for $i = 4$ and 5. If the relative area $a_i = 0.20$ for $i = 1$ to 5, consequently $f_2(\theta) = 0.60$.

If the element $A$ is part of the large grid and not the sum of finite elements, then the function $f_2(\theta_A)$ must be estimated using Eq. [2] by the mean value of the relative water content. Under the specified conditions of the area, it can be estimated as the simple mean of the value of each element; consequently $\theta_A = \bar{\theta} = 0.91$ and, in this case, the value of the function $f_2(\theta_A)$ is 0.40 and therefore $f_2(\theta_A) = f_2(\theta)$. As a result, under the given condition, a reduction in transpiration is observed from the fine to the coarse grid resolution. This simple example focused only on a single component of Eq. [1] for some particular conditions, but it helps to clarify that the evapotranspiration is influenced by the discretization process. The same holds true for all cases in which $f_2(\theta)$ has a nonlinear behavior, such as for dry and wet soils.

### Upscaling of Soil Properties

#### Soil Hydraulic Properties

An accurate determination of the soil hydraulic functions is a fundamental requirement for water and solute transport models. Soil water retention $\theta(b)$ and hydraulic conductivity $K(b)$ functions can be measured in the laboratory, directly in the field, or estimated from other soil properties using pedotransfer functions (PTFs). The soil hydraulic functions are usually given by a set of parameters for one of the common parameterizations. In this study, the van Genuchten–Mualem approach was chosen. Van Genuchten (1980) described the $\theta(b)$ and $K(b)$ relationship with six parameters: the van Genuchten parameters $\alpha$, $n$, and $m$, the saturated water content $\theta_s$, the residual water content $\theta_r$, and the saturated hydraulic conductivity $K_s$ (Vogel and Ippisch, 2008). Vogel et al. (1985, 2001) showed that the presence of the highly nonlinear $K(b)$ relationships near saturation can substantially impact the performance of the numerical solution of Richards’ equation in terms of the accuracy, stability, and rate of convergence of the invoked numerical scheme. Schaap and van Genuchten (2006) pointed out that the reduction in conductivity may become so dramatic that it may cause numerical instabilities in simulations of near-saturated infiltration when $1.0 < n < 1.3$.

In this study, the soil hydraulic parameters of the van Genuchten model were estimated from a soil texture data set (sand, clay, loam, humus, and skeleton content and porosity) using the PTF of Rawls and Brakensiek (1985). The resulting retention curve and saturated hydraulic conductivity have been corrected for the skeleton content by the PTF of Brakensiek and Rawls (1994).

#### Upscaling Soil Hydraulic Properties

The method used for upscaling soil hydraulic properties depends on several factors: data quality and data availability, the degree of soil heterogeneity, the boundary conditions, and the numerical scheme adapted (Renard et al., 2000; Diekkrüger, 2003; Vereecken et al., 2007).

Diekkrüger (2003) described different approaches to calculate the effective values: (i) the calculation of the mean parameters of the soil hydrologic properties, (ii) the ensemble retention $\theta(b)$ and conductivity $K(b)$ curves calculated by averaging the individual curves, (iii) calculation of the mean texture from the given textures and application of the PTF to the mean value, and (iv) calculation of the integral of the most important property used in the simulation.
and identification of a mean characteristic that fits the integral property. “Upscaling,” however, means to solve the relation

\[ \mu_{\text{eff}} = f (\omega, \mu_j) \]  

where \( \mu_{\text{eff}} \) is the effective parameter at the large scale, \( \mu_j \) are the parameters of \( j \) constitutive materials at the small scale, and \( \omega_j \) is their volume fraction. Samouelian et al. (2007) showed that the function \( f \) depends on the geometry and on the characteristics of the parameter that needs to be estimated. They used the power mean to estimate the effective hydraulic conductivity, which can be written for a general parameter as

\[ \mu_{\text{eff}} = \left[ \frac{1}{V} \int \mu (x)^p \, dV \right]^{1/p} \]  

where \( V \) is the volume of the domain, \( \mu \) is the local parameter, \( x \) is the spatial coordinate, and the power \( p \) reduces to the harmonic, geometric, and arithmetic means for \( p = -1, p = 1, \) and \( p \to 0 \), respectively. It was not our intention to investigate which of the available methods is most efficient. For the selected case study, we compared the effect of a simple method for upscaling soil properties with the effect that is imposed by coarsening the spatial resolution. We used the arithmetic mean for the determination of \( \alpha, \eta, m, \theta_s, \theta_r, \) and \( K_s \). For the parameters \( \alpha \) and \( K_s \), the logarithmic values were used to consider the lognormal probability density function.

### Evaluating Spatial Patterns of Relative Saturation

**Map Comparison**

With increases in computational power, distributed models have been more widely used to predict the spatially variability of hydrologic processes. Comparisons of lumped model output (e.g., discharge) are well known, but the comparison of spatial information (e.g., maps of soil moisture) requires a quantitative estimation of the agreement between patterns (Wealands et al., 2005). Comparison can be done between observed and simulated patterns or between patterns derived from different model configurations.

Comparing maps aids in the understanding and detection of temporal and spatial changes and in the analysis of results from different models, methodologies, or scenarios. It can be used in both the calibration and validation phases and, at a minimum, to analyze model uncertainty and sensitivity (Visser and de Nijs, 2006).

The main approaches used for pattern assessment are based on cell-by-cell error calculations. Wealands et al. (2005) presented an overview of available comparison methods. Techniques for comparing spatial maps can be grouped into four types: aggregate, cell-by-cell, fuzzy cell-by-cell, and higher order (Rose et al., 2009). In this study, map comparison techniques were applied to compare the relative soil saturation modeled from different soil and grid configurations. To reduce the amount of information, the mean relative saturation throughout the root zone was calculated for each element and compared between the different configurations.

The most common measure of similarity in the cell-by-cell calculation is the kappa statistic, which was introduced by Cohen (1960) for nominal maps. The kappa statistic is often used to assess the similarities between observed and predicted results. It is not only applied to geographical problems but is also used by many other disciplines such as the medical and social sciences. Hagen (2006) gave background information about the different disciplines where comparison methodologies have been investigated.

### Cell-by-Cell Map Comparison

The kappa value is a numerical measure of similarity between two maps based on a contingency table. It is determined by the percentage of agreement between two maps that have been corrected for the fraction of agreement that can be expected by random distribution (Hagen, 2002). When all values of the map are grouped in classes, the kappa statistic is computed based on an error matrix that counts how many cells were wrongly assigned to each category on both maps. The choice of the upper and lower limits of the classes influences the results, and maps that can appear similar to the human eye may be classified as being in poor agreement because similar values may be assigned for two different classes and the computed kappa statistic may be very small (Kuhnert et al., 2005). Recent extensions of the kappa statistic are the Kappa Location (KLoc) and the Kappa Histo (KHisto). These statistics are sensitive to the respective differences in the location and the histogram shape for all categories. For more details concerning the Kappa Location and Kappa Histo, see Hagen (2002). Kappa, KLoc, and KHisto are connected through the multiplicative relation

\[ \text{kappa} = \text{KLoc} \times \text{KHisto} \]  

Thus, given a fixed value for kappa, KLoc will increase if KHisto decreases and vice versa. When cells at the same location have the same values in two different maps, the KLoc = 1.0. When two maps have the same number of elements for each category, then KHisto = 1.0. The strength of the agreement was classified by Landis and Koch (1977) as described in Table 1. The measure can comprise values between −1 and 1, where 1 is obtained in case

<table>
<thead>
<tr>
<th>Kappa statistic</th>
<th>Strength of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.00</td>
<td>poor</td>
</tr>
<tr>
<td>0.00–0.20</td>
<td>slight</td>
</tr>
<tr>
<td>0.21–0.40</td>
<td>fair</td>
</tr>
<tr>
<td>0.41–0.60</td>
<td>moderate</td>
</tr>
<tr>
<td>0.61–0.80</td>
<td>substantial</td>
</tr>
<tr>
<td>0.81–1.00</td>
<td>almost perfect</td>
</tr>
</tbody>
</table>

Table 1. Definition of level of agreement used to describe the kappa statistic (Landis and Koch, 1977).
of perfect agreement, and zero or negative values are obtained in cases where the value is equal to or less than what is expected in a random distribution. It should be kept in mind that the kappa values strongly depend on the definition of the classes.

**Application of Hydrogeosphere to a Small Headwater Catchment**

**Site Characterization**

The focus of this study was the forested headwater catchment Wüstebach (Fig. 2), which is a tributary of the Erkensruhr River in Germany with a size of about 27 ha (0.27 km²). The Wüstebach basin has a warm temperate oceanic climate. The mean annual temperature is 7°C. The mean sunshine during the year is 1500 to 1600 h, while the mean annual precipitation usually ranges from 1100 to 1200 mm. In the upper Wüstebach basin, a snow cover deeper than 10 cm is observed during 20 to 30 d yr⁻¹, which was not considered in the actual version of the model. The wind comes predominantly from the western direction, and during 90% of the year the velocity is >1.5 m s⁻¹ (Minister für Umwelt, 1989). The mean elevation is 595 m above sea level and the catchment has a slope of 9%.

**Soil Characterization**

Soil characterization was made by the Geologischer Dienst Nordrhein-Westfalen (Geological Survey of North-Rhine-Westphalia) at a scale of 1:2500, in which the surface of the catchment was assembled into nine soil types (see Fig. 2) that were divided into 61 soil mapping units. Each of these mapping units was divided into three soil horizons, for a total of 183 soil layers. Most of the soil units in the map have the same properties but are accounted as separate units because they are located in different places. There are 31 different soil textures present in the catchment, for which the sand, clay, silt, humus, and rock fragments or skeleton (>2 mm) content are available (Richter, 2008). Soil hydraulic properties were estimated as described above and are given in Table 2.

**Surface Roughness**

The soil microtopography determines several processes that are involved in surface water dynamics. The surface processes can be principally regulated in the model through the Manning coefficients (in the x and y directions), rill storage, and obstruction storage. They impact the surface runoff initiation and water flow, and they can substantially change the shape of the hydrograph (Woolhiser et al., 1997).

Spatially distributed models need an estimation of the roughness plan. In densely vegetated floodplains, the major roughness is caused by trees and brush (Arcement and Schneider, 1989). The rill storage represents the amount of water storage that must be reached before flow starts, and the obstruction storage is a term used in forest or agricultural settings to represent the resistance of the vegetation to the flow (Panday and Huyakorn, 2004).

In the Wüstebach catchment, a large roughness was observed, e.g., due to exposed roots, high tree density, and downed trees. Manning’s coefficient, as a friction factor, has an effect on the shape of the hydrograph because low Manning coefficients produce high peaks in the hydrograph, while on the contrary, high values reduce the peaks. The water balance didn’t change in our case because the soils are highly permeable, and surface runoff is of minor importance. In the absence of their direct measurement in this study, a Manning’s coefficient of 0.65 s m⁻¹/³, a rill storage height of 0.25 m, and an obstruction storage height of 0.25 m were manually chosen based on the literature (Arcement and Schneider, 1989; Therrien et al., 2007); because the cover of the catchment was homogeneous, these parameters were assumed to be homogeneous for the whole catchment. The validity of the chosen parameters was also supported by in situ inspection with regard to their physical meaning.

**Vegetation Parameters**

The catchment is densely forested by Norway spruce [Picea abies (L.) H. Karst.], a species characterized by a shallow root system; the plant coverage is about 90%. This location is also rich in low vegetation, like plants that grow in water-saturated areas [e.g.,

---

Fig. 2. Location and soil map of the Wüstebach catchment with the soil types as provided by the Geological Survey of North-Rhine-Westphalia.
Sphagnum spp. and Cirsium palustre (L.) Scop. (J. Deckers, personal communication, 2009).

To correctly compute the evapotranspiration, model parameters have to be adapted to the given situation. Root depth was estimated directly in the field. As the catchment is relatively wet due to high precipitation, the dry part of the function $f_2(\theta)$ has a particular impact on the results only during extremely dry years. Because the soil types are similar in texture (see Table 2) and to simplify model analysis, the relative saturation of the water content at the wilting point and field capacity were set as being homogeneous for the whole domain. Other evapotranspiration parameters were taken from the literature (e.g., Dickinson et al., 1986; Breuer et al., 2003; Panday and Huyakorn, 2004; Li et al., 2008; Gigante et al., 2009). In particular, all parameters that have been used for the simulation of evapotranspiration are reported in Table 3.

### Atmospheric Forcing

The analysis was performed for the year 2007. The atmospheric forcing was taken from the COSMO-DE analysis, which is derived from the COSMO weather forecast (Steppeler et al., 2003) and corrected by measurements. The total amount of precipitation for 2007 was 1468 mm, which was a wet year compared with the mean.

To compute the reference evapotranspiration $ET_0$ (mm d$^{-1}$) the guidelines of Allen et al. (1998) were used, which compute reference crop evapotranspiration at hourly time steps using the Penman–Monteith equation. The cumulative potential evapotranspiration computed for 2007 was 978 mm yr$^{-1}$.

### Discretization

The first step in applying HydroGeoSphere is grid generation, for which the finite elements preprocessor GRID BUILDER (McLaren, 2005) was used. The finite element grid is generated.
automatically in the two-dimensional plane and is projected in the vertical to form a three-dimensional grid. A 10-m digital elevation model was used to define a two-dimensional, triangular-element mesh representing the top of the model domain.

Two different horizontal grid configurations were compared: Case A, which had element lengths of 25 m and fine refinement at the river zone with element lengths <10 m, and Case B with an overall resolution of 100 m. For Case A, a grid refinement was performed to allow a more realistic description of the local phenomenon at the river zone. Case B had no refinement at the river zone to verify the ability of the model to solve the water flux problems for larger scales. The river under investigation is a headwater catchment and usually this zone is not refined when the small catchment is part of a larger area. The grid configurations in the two-dimensional horizontal domain are shown in Fig. 3. In both configurations, the vertical domain was divided into 23 layers, with differing layer thicknesses: the upper 15 cm was divided into six layers, each with a thickness of 5 cm; and the underlying 1 m of soil was divided into 10 layers, each with a thickness of 10 cm. From the soil map it can be concluded that the mean soil cover is 1.50 m thick, overlying an impermeable bedrock.

In Case A, the soil hydraulic parameters were taken directly from the soil map with, in fact, only a few elements having more than one soil type. Therefore, the dominant soil type in the element was chosen as being representative. The domain generated by the grid in Case A was divided into 324 homogeneous elements. The upper 10 cm was considered to be homogeneous across the whole catchment because this is the organic litter layer that has high permeability and high water holding capacity. The elements of the domain generated from the grid in Case B had more than one soil type, some elements consisting of >12 soil textures. Therefore, the vertical domain was assembled into five horizons of 0.10, 0.20, 0.30, 0.40, and 0.50 m from the top to the bottom. For each of the 108 elements and for each soil horizon, the soil properties were estimated using the mean averaging method as given in Eq. [5].

The third simulation, Case C, had the same three-dimensional grid configuration as Case A, but the soil configuration was the same as Case B. Therefore, Cases B and C had the same spatial variability of soil parameters but with a different spatial discretization, and Cases A and C had the same spatial discretization but different soil properties. With this setup, it is possible to differentiate between the effects due to upscaling soil properties and due to grid resolution.

**Model Setup**

To couple the surface and the subsurface flow domains, HydroGeoSphere uses a first-order exchange coefficient (Ebel et al., 2009). At a given location of the top domain, the model has two nodes, one in the surface domain and one in the subsurface. The leakage factor, defined in HydroGeoSphere as a coupling length between the surface and the subsurface domain, represents the resistance to the exchange between nodes in the surface and nodes in the porous medium (Therrien et al., 2007). A low value of the coupling length coefficient represents rapid exchange of flow

Table 3. General parameters of the evapotranspiration model used for all simulations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaporation depth, m</td>
<td>0.2</td>
<td>estimated</td>
</tr>
<tr>
<td>Root depth, m</td>
<td>0.5</td>
<td>measured</td>
</tr>
<tr>
<td>Leaf area index, m⁻²</td>
<td>6.7</td>
<td>Breuer et al. (2003)</td>
</tr>
<tr>
<td>Relative saturation at field capacity, m⁻³</td>
<td>0.4</td>
<td>estimated</td>
</tr>
<tr>
<td>Relative saturation at wilting point, m⁻³</td>
<td>0.3</td>
<td>estimated</td>
</tr>
<tr>
<td>Relative saturation at oxic limit, m⁻³</td>
<td>0.85</td>
<td>estimated</td>
</tr>
<tr>
<td>Relative saturation at anoxic limit, m⁻³</td>
<td>0.97</td>
<td>estimated</td>
</tr>
<tr>
<td>Energy limiting stage saturation, m⁻³</td>
<td>0.4</td>
<td>Therrien et al. (2007)</td>
</tr>
<tr>
<td>Evaporation limiting stage saturation, m⁻³</td>
<td>0.3</td>
<td>Therrien et al. (2007)</td>
</tr>
<tr>
<td>Canopy storage parameter per leaf area index, m</td>
<td>0.0002</td>
<td>Dickinson et al. (1986)</td>
</tr>
<tr>
<td>Initial interception storage, m</td>
<td>0</td>
<td>assumed</td>
</tr>
</tbody>
</table>
between the two domains, and conversely, a high value represents slow exchange. We tested a range of these values with regard to the total water balance. For this purpose, the sensitivity to the leakance factor was very low, but the simulation duration changed significantly because low leakance factor values require more computation time. A value of 0.1 m was chosen for the coupling length, which is a compromise between the maximum computational efficiency and numerical stability and the plausibility of the modeled process. The boundary conditions were expressed in terms of flux; in particular, rainfall and potential evapotranspiration, at hourly time steps, were applied at the surface and no-flow boundary conditions were imposed for the bottom and lateral domains.

The catchment is relatively wet because of high precipitation amounts; therefore, the initial conditions for the first simulation were set to saturation at time \( t = 0 \). The first simulation was considered a spin-up to compute realistic initial conditions that were consistent with the climate and soil conditions. The required times for the spin-up processes were similar to the duration of the simulation, as reported in Table 4. All analyses reported below were performed using 1 yr of spin-up, which was proven to be enough to provide stable initial conditions.

HydroGeoSphere is not parallel processing, and the simulations were performed on a processor core with 1.81 GHz and 1.64 Gb of RAM.

**Results and Discussion**

**Hydrograph Analysis**

The first step in evaluating the effect of upscaling is analysis of the hydrographs. The analysis was performed by comparing the obtained output from the different Cases A, B, and C with each other and with the observations.

The most realistic simulation is expected to be the one with the best spatial resolution and the original soil map. Therefore, the “benchmark output” is the configuration of Case A. In Fig. 4, the observed and simulated discharges for Cases A and B are given. It is noteworthy that the soil properties were not calibrated but derived from the soil map using PTFs; furthermore, vegetation parameters were chosen from the literature (Therrien et al., 2007). The computer time required restricted automatic calibration, but further calibration of the vegetation parameters might improve the performance of the model.

As can be seen in Fig. 4, discharge was partly overestimated in June and underestimated from October to November, but the general features of the measured discharge were captured by the simulations. Considering that the soil hydraulic properties were not calibrated, the hydrograph for Case A shows good agreement with the hydrograph for the observations, and they have a high Pearson’s correlation coefficient \( r = 0.83 \). The total discharge observed for 2007 at the hydrograph station was 827 mm yr\(^{-1}\). As given in Table 4, the total simulated discharge was 799 mm yr\(^{-1}\), the interception was 188 mm yr\(^{-1}\), and the actual evapotranspiration was 468 mm yr\(^{-1}\) for Case A. The hydrograph for Case B shows moderate agreement with the observations \( r = 0.68 \).

Even if the \( r \) value for Case A was higher, however, it would be difficult to decide which of the hydrographs best fits the observed values. In addition, the lack of actual evapotranspiration measurements hampers our ability to decide which value for the estimated actual evapotranspiration is the correct one, although a clue would be that the total discharge was simulated with a sufficient accuracy only in Case A. The mean water balance error was lower for the coarser discretization because of the smaller lateral gradients.

It is quite clear that a large difference (\( \sim 100 \text{ mm yr}^{-1} \)) between the two hydrographs is linked to a similar difference in the actual

![Fig. 4. Comparison of the simulated (Case A with grid element lengths of 25 m and Case B with element lengths of 100 m) and the observed discharge values.](image-url)
evapotranspiration. Case B shows some very high peaks in the discharge and Case A has very low values in the dry period (March–April). From this comparison alone it is not possible to identify the reason for these differences. Consequently, the analysis must be supported by the results from Case C. In fact, Cases C and A have similar values for all the components of the water balance and also for the correlation coefficient. This indicates that the increase in the discharge is not because of the method used for upscaling the soil hydraulic parameters. The reduction in evapotranspiration for Case B can be explained by effects discussed previously concerning the influence of the spatial grid resolution.

The difference in the interception values between Cases A and C (see Table 4) can be explained as a result of the adaptive time step scheme used in the simulation. The canopy interception process was simulated as if it were bucket storage; consequently, the water that had not evaporated in one time step was used as the initial storage in the next time step. Depending on the internal time stepping, the water interception in the two simulations may differ, even though all other parameters are constant.

To evaluate the model performance, the quality measures $r$ and the root mean squared error (RMSE) between the observed and simulated values at an hourly time step were computed. These values are reported in Table 4 and are accompanied by the water balance components for the three cases.

Two parameters that are difficult to estimate are the oxic and anoxic limits, which are used in Eq. [2]. To evaluate the differences in simulation behavior, the sensitivity of these parameters had to be investigated. Therefore, different simulations were performed using the setup of Case B. The results (Fig. 5) show that the anoxic and oxic limits have a significant effect on the water balance and, therefore, should be determined with care. Actual evapotranspiration varies between 300 and 600 mm yr$^{-1}$, which make these parameters some of the most sensitive parameters out of the entire simulation model.

**Analysis of the Spatial Discretization Effects**

The differences in the water balance, as shown in Table 4, can be caused by several factors. Coarsening the resolution results in a loss of topographic information because of the smoothing of the soil surface. Furthermore, the smaller lateral gradients can cause different numerical errors to occur. Therefore, the differences between the simulations do not have an obvious or simple cause and effect relationship. To determine whether the differences could be explained by numerical phenomena, the same configurations were used, except with the actual evapotranspiration computed as a linear function of the soil moisture. To avoid the effect highlighted above and to simulate a linear relationship between the soil moisture and actual evapotranspiration as given in Eq. [2], the soil moisture content at the wilting point ($\theta_{wp}$) was fixed to zero and to one at field capacity and the oxic and anoxic limits ($\theta_{oc}$, $\theta_{o}$, and $\theta_{an}$, respectively). In this numerical experiment, the pure effect of the grid size on the numerical accuracy could be determined. The water balance for these configurations is given in Table 5. Although the soil moisture contents of the coarse grid might not be equal to the average of the fine grid, the differences between the hydrologic components were smaller than discussed above and given in Table 4. In particular, the difference of approximately 100 mm in the actual evapotranspiration between Case B and Cases C and A reduced to ~20 mm when the linear relationship was used.

In addition, the estimation of a relative error between two grids might help to analyze the sources of these differences. The relative error ($\epsilon$) is determined as (Celik et al., 2008, Roache, 1994)

$$\epsilon_{ji}^k (\%) = \left| \frac{g_i - g_j}{g_i} \right|$$

**Table 5. Water balance and model performance for the model configuration with the linear relationship between soil moisture and actual evapotranspiration.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interception, mm yr$^{-1}$</td>
<td>154</td>
<td>159</td>
<td>149</td>
</tr>
<tr>
<td>Actual evapotranspiration, mm yr$^{-1}$</td>
<td>590</td>
<td>606</td>
<td>586</td>
</tr>
<tr>
<td>Discharge, mm yr$^{-1}$</td>
<td>717</td>
<td>698</td>
<td>723</td>
</tr>
<tr>
<td>Mean water balance error, mm h$^{-1}$</td>
<td>$4.50 \times 10^{-4}$</td>
<td>$2.14 \times 10^{-4}$</td>
<td>$4.17 \times 10^{-4}$</td>
</tr>
<tr>
<td>CPU time, s</td>
<td>210,500</td>
<td>5650</td>
<td>195,000</td>
</tr>
</tbody>
</table>
where $g_i$ and $g_j$ are the global variables (i.e., discharge and actual evapotranspiration) for the fine grid $i$ and the coarse grid $j$, respectively, and the superscript $k = n$ or $l$ is for the cases with a nonlinear or the linear relationship, respectively, between soil moisture and actual evapotranspiration. As can be seen in Table 6, the relative error between the fine and coarse grid configurations changed substantially from the nonlinear to the linear relationship. Using the linear relationship, the differences are on the order of 3% instead of the 20% obtained using the nonlinear relationship. It is not possible to differentiate this 3% into numerical error, error generated by the upscaling soil hydraulic properties, and error generated by coarsening the discretization, but it can be concluded that the effect of the spatial discretization on the evapotranspiration process was dominant.

### Pattern Comparison

It is not sufficient to just evaluate the quality metrics of the model performance, such as the coefficient of correlation and the RMSE (as reported in Table 4). Even if they give a first impression about the model performance, an evaluation of the spatial patterns of the soil water dynamics is required for distributed modeling.

In this study, the spatial variability of the patterns was represented by the kappa statistic. For the three model configurations, the comparison of the patterns of the relative saturation (as a percentage of the volume) was made using the Map Comparison Kit (Hagen et al., 2006). To compare maps of different resolutions, they must be transformed to a fine grid (here, 2 by 2 m), which was done using ArcGIS software (ESRI, Redlands, CA). A semiautomated procedure does not allow rapid utilization of all the available information; therefore, the analysis was performed at 5-d intervals and for the mean relative saturation of the root zone only (the upper 50 cm). In particular, the comparison between Cases A and B (A–B), A and C (A–C), and finally B and C (B–C) were made for 71 soil moisture snapshots during the year, which resulted in one comparison every 5 d.

Figure 6 shows the maps of relative saturation for Cases A, B, and C and the result of the comparison among them for 10 February, for which the kappa values showed good agreement. In this figure, it is clear that the coarser the resolution, the larger the wet area is compared with the fine resolution, which is the main difference between these two maps. According to Table 1, the A–C comparison can be classified as “almost perfect agreement.”

In Fig. 7, the maps for 1 May are shown, for which the kappa values showed poor agreement. Low kappa values are often present during the drying period after strong rainfall events resulting

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Relative error</th>
<th>Linear relationship</th>
<th>Nonlinear relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A vs. B</td>
<td>C vs. B</td>
<td>A vs. C</td>
</tr>
<tr>
<td>Discharge</td>
<td>2.7</td>
<td>3.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Actual evapotranspiration</td>
<td>2.5</td>
<td>3.4</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 6. Relative error between the two grid configurations for the cases with linear and nonlinear relationships between soil moisture and actual evapotranspiration. Cases A and C had a grid element length of 25 m and Case B had element lengths of 100 m; Case A used heterogeneous soil properties, while Cases B and C used upscaled soil properties.

![Fig. 6. Maps of the relative saturation (upper) for Cases A, B, and C and the result of the comparison among them (lower) for a time step where the kappa values indicated good agreement (10 February). Cases A and C had a grid element length of 25 m and Case B had element lengths of 100 m; Case A used soil properties, while Cases B and C used upscaled soil properties.](image)

![Fig. 7. Map of the relative saturation (upper) for Cases A, B, and C and the result of the comparison among them (lower) for a time step where the kappa values indicated poor agreement (1 May). Cases A and C had a grid element length of 25 m and Case B had element lengths of 100 m; Case A used soil properties, while Cases B and C used upscaled soil properties.](image)
in a high peak discharge. As shown in Fig. 6 and 7, low kappa values were observed near the river, while high values were computed for the region distant from the river. This result is related to the different spacing of computational nodes because in large grids, the variability of soil moisture decreases due to the effects of averaging (Western et al., 2002).

In Fig. 8 to 10, the results for the simulation period are given as time series and box plots. The kappa values increase in the rising limbs of the hydrograph and decrease in the falling limbs. This was observed, for example, between the beginning of February and the end of April, where the discharge started at low values and increased with time.

The interpretation of Fig. 8 to 10 supports the discussions above, which argued that Cases A and C are similar. The mean kappa values are classified as having “substantial agreement,” which becomes an “almost perfect agreement” with regard to KLoc. The mean kappa for the comparison A–B shows a “fair agreement,” similar to the mean value for the comparison B–C, though the latter has better performance measures than A–B.

In Table 7, the mean values of kappa, KLoc, KHist, and the correlation coefficient between the simulated discharges are reported for each comparison.

**Summary and Conclusions**

An integrated surface–subsurface, three-dimensional, finite element model was applied to a forested headwater catchment to evaluate the influence of soil heterogeneity and spatial discretization on the water balance. The distributed hydrologic model HydroGeoSphere was applied to the 27-ha Wüstebach catchment for a simulation period of 1 yr at hourly time steps. The analysis of...
the water balance was supplemented by a quantitative assessment of the effects on the spatial patterns in soil moisture.

Different cases were simulated with different spatial discretizations and with upscaled soil hydraulic parameters. Analysis of the hydrographs and the total amount of the actual evapotranspiration showed that the volume averaging method, used for the estimation of the soil hydraulic properties, gave reasonable results. Contrary to the upsampling of the soil hydraulic properties, the spatial discretization had a large effect on the entire water balance. In particular, a large grid produced more discharge and less actual evapotranspiration than a small grid. This can be explained by the nonlinear relationship between water stress and soil moisture.

In addition to analyzing the discharge, the spatial patterns of soil moisture were investigated using kappa statistics. The analysis confirmed that the results obtained by upsampling the soil hydraulic properties were sufficiently similar to the cases with completely heterogeneous parameters. This highlights the fact that the resolution of the spatial discretization has a larger impact on the soil moisture pattern than the aggregation of the soil properties. Maps of relative saturation using the same spatial discretization showed nearly perfect agreement, while maps with differing spatial discretization showed only fair to moderate agreement. By using a linear relationship between water stress and evapotranspiration, we can show that during the upsampling process, evapotranspiration was the main source of error caused by coarsening the spatial resolution. New methods for upsampling parameters used in evapotranspiration models are required to overcome this limitation.

It is noteworthy that this study was performed for a real catchment, that the spatial configurations were set at 25 and 100 m, and that the model parameterization was implemented without calibration. Although no general rules can be given concerning the optimal grid size, it is always recommended to use the finest grid configuration balanced with the availability of data and the required computational time to guarantee an accurate solution. A stepwise reduction of the grid size may give some hints on the critical values that should not be exceeded. This study, however, demonstrates the limitation of one of the most frequently used evapotranspiration approaches. Future research should improve the methods for upsampling vegetation-related parameters because they are highly sensitive. The upsampling of soil hydraulic properties has been investigated for many years and sophisticated methods are available. Nowadays, the focus lies on a more sophisticated description of the evapotranspiration process because this is often modeled in a very rough and empirical way, which leads to large uncertainties in discharge dynamics and in the spatial patterns of soil moisture.

Acknowledgments
We thank Rob McLaren of the University of Waterloo for allowing us to use his software package (HGS and Grid Builder), Annika Schomburg for processing the COSMO-DE analyses, which were provided by the DWD (Deutscher Wetterdienst, German Meteorological Service), and the Deutsche Forschungsgemeinschaft (DFG) for financial support of sub-project C1 of the Transregional Collaborative Research Centre 32 “Patterns in Soil–Vegetation–Atmosphere Systems.”


