Regionalization in Hydrology (Proceedings of a conference held at Braunschweig, March 1997). IAHS Publ. no. 254, 1999.

Regionalization concept for the prediction of large-scale water fluxes

HELGE BORMANN, BERND DIEKKRÜGER, CHRISTIAN RENSCHLER

Geographische Institute der Universität Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany

e-mail: h.bormann@uni-bonn.de

OTTO RICHTER

TU Braunschweig, Institut für Geographie und Geoökologie, Langer Kamp 19c, D-38106 Braunschweig, Germany

Abstract A combined model and regionalization concept is presented which enables the user to simulate water fluxes at a wide range of scales (from sites to regions) without changing the model concept. The necessary aggregation is achieved by defining hydrologically similar areas. The similarity of the hydropedotopes is defined by a cluster analysis of simulation results (e.g. monthly values of actual evapotranspiration, groundwater recharge, surface runoff and interflow). The results of an application of this concept are presented. Water fluxes from a 1000 km^2 sized catchment were calculated using scale dependent information density.

INTRODUCTION

There is wide concern about the scale dependant prediction of water and matter fluxes. The model systems which have been developed in the last decades are investigated to assess their applicability on different scales and in catchments with varying properties. The spatial autocorrelation of soil properties and water fluxes can be regarded at small scales but not when investigating large catchments (Blöschl, 1996). Natural structures at the small scale can be represented by model parameters. This cannot be done in large-scale modelling because the spatial discretization cannot be as fine as it is on small scales due to decreasing knowledge.

So when shifting to a coarser scale it is important to derive effective or representative parameters. In this case the concepts of "hydrological response units" (HRU) or "representative elementary areas" (REA) are used predominantly (Wood, 1995). The aim is to determine hydrologically similar areas. In this case the landscape consists of numerous HRUs or REAs. However questions arise as to how to define similarity and how to derive model parameters from catchment properties. Aggregating parameters is not feasible due to nonlinearities in the model equations. One possibility is to define similarity by simulation results. Alternatively the model system chosen can be simplified. Defining HRUs mostly leads to a change in the chosen model concept, often from distributed to lumped models. The choice of the model depends not only on the database but also on the object of the investigations. Lumped models are not suitable for the prediction of the spatial distribution of evapotranspiration (Wood,

1995) and cannot be used to consider the small-scale variability in runoff production. Blöschl (1996) stated that there are a number of regionalization concepts, and each has validity.

The special research programme of the Deutsche Forschungsgemeinschaft (DFG) "Regionalization in Hydrology" deals with the essential processes of the water cycle of the land surface: soil water flux, evapotranspiration, percolation, interflow and surface runoff. These hydrological processes are described by a physically based model system at the local scale. The chosen model calculates one-dimensional water fluxes with a high spatial and temporal resolution. Therefore the Richards' equation—which is a partial differential equation—has to be solved numerically for the soil water flux. The computational effort is high and therefore it is not feasible to solve the Richards' equation for numerous hydro-pedotopes in large catchments. In this study the aggregation necessary is achieved by using representative parameters but not by a change of the model concept.

The main goal of the project is to develop a model and regionalization concept which enables the simulation of water fluxes at a wide range of scales (from sites to regions) without a change in model concept. The strategy presented is limited, as are the other concepts. The development of a regionalization concept and its evaluation are therefore both central aspects of interest. No additional field measurements were to be performed within the project. Model parameters were to be derived from generally available information such as soil maps, geological maps, digital elevation models and weather data. A model calibration was neither feasible nor required because of the chosen distributed approach. For the development of the regionalization concept a database for the north of Germany was available. Two research catchments (85 km north and 35 km south of Braunschweig) were investigated for many years in another DFG research project: Water and Matter Dynamics in Agroecosystems. The applicability of the regionalization concept was tested in the "obere Leine" target catchment which is situated in the northern German Midlands and has an area of about 1000 km^2.

MODEL CONCEPT

In this study the physically based, one-dimensional model system SIMULAT (Modellsystem zur Berechnung der Wasser- und Stoffdynamik landwirtschaftlich genutzter Standorte; Diekkrüger, 1996) is applied. The soil water flux at the local scale is described by the Richards' equation. Infiltration is also calculated by the Richards' equation using a semi-analytical solution (Smith & Parlange, 1978). The calculation of the potential evapotranspiration is accomplished by the Penman-Monteith equation. Additional modules for the computation of interception, baseflow and snow cover are included. A complete description of the model is given in Diekkrüger (1996).

The first step towards a simulation of the water fluxes in large drainage areas is to derive homogeneous areas. This is achieved by an overlay of different information layers by means of the GIS ARC/INFO. The smallest homogeneous area (ecotope) was created by overlaying of soil, geological, land use, topographic and climate data.

To determine the discharge curve at the catchment outlet the calculated water fluxes of each ecotope are routed to the catchment outlet. For large catchments the application of the St Venant equation is too complicated and needs too much effort.

Therefore a simplified flood routing approach is used which is available in ARC/INFO. From the mean ecotope size (6 ha) and the mean drainage network density (1.5 km km^{-2}) it can be shown that each ecotope is connected with the drainage network. Thus the ecotopes can also be assumed to be unaffected by lateral flow components. A spatial distribution of flow velocities was derived from slope and land use.

REGIONALIZATION CONCEPT

For the simulation of large drainage areas various alternative concepts might be considered.

(a) **A catchment consists of a large number of homogeneous hydro-pedotopes, and water flux is simulated for every unit.** Deriving the spatial distribution of ecotopes and simulating the water fluxes using the SIMULAT model system principally enables the user to predict the temporal and spatial distribution of water fluxes in a given catchment. The size of the ecotopes then depends on the database available. But the number of derived ecotopes increases rapidly with the size of the catchment. Although computer power is increasing very fast it is unrealistic to simulate all derived ecotopes for regional model applications due to available computer time.

(b) **Use of a database with a coarser spatial resolution to simulate the water fluxes of the ecotopes resulting from the overlay operation**. Simulating the water fluxes based on a coarser database involves the use of unknown aggregation rules. The data are not usually aggregated for hydrological purposes. Due to this aggregation process significant errors and trends may occur (Bormann *et al.*, 1996; Bormann *et al.*, 1999). Alternatively it is possible to scale up parameters from a highly resolved database to derive effective parameters. An effective parameter replaces an ensemble of realizations in a catchment by one single value. An effective model on the global scale results from the calculation of the expected values of the local variables of the ensemble of possible realizations of stochastic processes. Contrary to groundwater systems based on the Darcy equation, the application of mean model parameters is not feasible due to the nonlinear model behaviour (Blöschl, 1995).

(c) **Use of a simplified model for the simulation of water fluxes for large catchments**. Each scale can be characterized by its dominant processes. So the scale-dependent hydrological behaviour (in space and time!) can be simulated by neglecting processes which are only important at smaller scales. But the question arises as to how to realize this scale dependent change in the model concept. It is necessary to ascertain which processes are dominant at the different scales under distinct conditions. Applying the regionalization approach, a scale invariant process structure was assumed with the same processes acting at different scales.

(d) **Sort all available hydro-pedotopes into groups of similar hydrological behaviour and simulate only one member of each group**. Averaging model parameters does not lead to reasonable simulation results. Thus the regionalization concept presented tries to group ecotopes showing similar hydrological behaviour. The classification is based on a cluster analysis. The furthest neighbour method based on the Euclidean distance (Nieschulz, 1997) as well as the Ward method based on the squared Euclidean distance showed reasonable results.

Monthly water fluxes (actual evapotranspiration, groundwater recharge, interflow and surface runoff) are used as input data for the cluster analysis to consider the seasonal dynamics of the vegetation cover. As the number of simulated ecotopes is reduced the cluster analysis cannot be based on the simulation results of all "real" ecotopes in a given drainage area, derived by overlaying digital information of soil properties, land use and topography. Thus "theoretical" ecotopes are defined. Therefore homogeneous soil profiles are chosen which cover the entire soil texture triangle (all soil texture classes). All combinations of soil profiles with different crops as well as slopes and different groundwater tables (= theoretical ecotopes) are simulated once.

Applying the cluster analysis, hydrologically similar ecotopes are grouped. Using the definition of similarity only one representative ecotope has to be used for each cluster in further simulations. The main problem of the regionalization concept is to link the real ecotopes, derived from catchment properties, to the theoretical ecotopes based on simplified soil profiles. Depending on the database the real soil profile may consist of a number of horizons of different thickness. Each horizon is characterized by a set of soil properties (e.g. texture, density, organic matter content). Neither mean model parameters nor mean soil properties result in meaningful parameters for the whole soil profile. So integral measures are needed which summarize the complex soil properties within one single value. The combination of a few integral measures, called regionalization indices, allows a definite assignment of real soil profiles to theoretical soil profiles.

In a first approach (Nieschulz, 1997) the "water available for plants" was chosen based on the knowledge of soil controlled water fluxes. This first regionalization index is now supplemented by two other indices, the "travel time" and the "height of capillary flux" (Finke, 1992). The travel time describes the time a water molecule needs to pass through the rooting zone assuming a non-dispersive vertical water flux. The height of capillary flux is a measure of the capillary rise into the rooting zone. Combining the regionalization indices height of capillary flux and water available for plants results in a suitable measure for the "total amount of water available for plants". In addition to the regionalization indices used in soil science, Nieschulz (1997) proposed indices based on mathematics (eigen values of the matrix of soil properties as texture, porosity, etc.). Several methods are suitable for the determination of the representative of each cluster. Based on the definition of similarity all members in a cluster are similar in their soil hydrological behaviour. If the classification is detailed enough the choice of the representative ecotope is of minor importance. The selected ecotope is then called the representative ecotope. It shows the mean hydrological behaviour but does not rely on effective parameters in the usual sense. This regionalization concept is described schematically in Fig. 1.

The ensemble of theoretical ecotopes has to be classified only once. Afterwards the regionalization indices are determined for all ecotopes in the catchment of interest. The indices are statistical quantities and so the computational effort is very small. After the assignment of the real ecotopes to the clusters based on theoretical ecotopes, the water balance of the study catchment can be calculated using only a few representative ecotopes. By applying this concept, no temporal and spatial aggregation

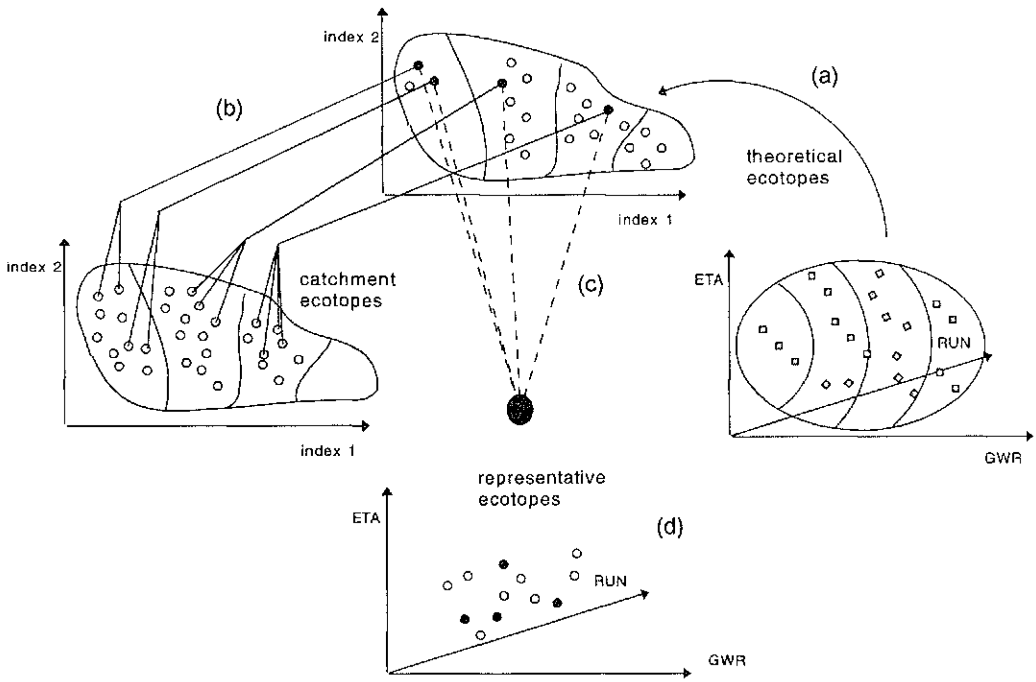

Fig. 1 Regionalization scheme (a) simulation and classification of "theoretical" ecotopes, definition of regionalization indices; (b) determination of regionalization indices of all "real" ecotopes, assignment to the clusters of "theoretical" ecotopes; (c) choice of representative ecotopes; and (d) simulation of the water fluxes of the representative ecotopes for the calculation of the regional water balance.

is performed. The spatial pattern which results from the GIS analysis is preserved as well as the temporal discretization of simulation results. A more detailed description is given by Nieschulz (1997), Lücke (1997) and Diekkrüger *et al.* (1998).

Lücke (1997) showed that the results of the cluster analysis concerning surface runoff production are significantly influenced by initial and boundary conditions. For different rainfall events different classifications are possible. Effective parameters cannot be determined and the surface runoff cannot be calculated for a single simulation as surface runoff is not produced until the precipitation rate exceeds the infiltration rate. In this case the minimum infiltration rate equals the saturated water conductivity (k_s).

A sensitivity analysis showed that the runoff production depends critically on the variability in k_s (Bormann, 1999). Other model parameters are only of minor importance. The frequency distribution of k_s can be taken into account by replacing a slope by a number of parallel and homogeneous flow strips which vary in k_s (Smith *et al.*, 1990). Therefore both the expected value and coefficient of variation of k_s is needed.

INFORMATION HIERARCHY

As the size of a catchment increases the information density decreases. This results in a deterioration of the spatial and temporal resolution of input data. The question arises as to whether for regional investigations the decreased information density is sufficient when using physically based models. According to Bormann *et al.* (1999) the spatial pattern of the water fluxes (e.g. actual evapotranspiration) in a given catchment shows significant differences in resolution and structure when reducing the resolution of the used soil map. At some sites annual differences of up to 15% were calculated.

Crucial deviations were observed when publicly available data from the Deutscher Wetterdienst (German Meteorological Service) were used instead of our own weather measurements. The aspects of temporal resolution were investigated first. The information loss caused by a coarse temporal resolution of weather data (e.g. daily minimum, mean and maximum values instead of an hourly resolution) is compensated for when using disaggregation rules. Errors based on the disaggregation process are negligible for evapotranspiration and groundwater recharge, but not for surface runoff. The aspect of spatial transference of weather data is of major importance. When calculating the actual evapotranspiration, the small-scale variability of air humidity and precipitation in particular are decisive. Errors caused by transferring rainfall data can reach 20% of the annual evapotranspiration. For further information see Bormann *et al.* (1996). Investigations concerning the quality of the digital elevation model (DEM) show the influence of relief structure for the discharge curve. So a grid with a high vertical resolution should be used (Thieken *et al.*, 1997).

DATABASE

In contrast to the test catchments, the density of measurement stations in the target catchment is low. Only two meteorological stations and 34 raingauges are available, providing data for differing time periods. The monthly averages of the precipitation stations were grouped into six regions using a cluster analysis. The regions which show a similar precipitation regime were constructed by Thiessen polygons and for each region a representative station was selected. Then the daily values of the six representative stations were disaggregated to hourly values using a stochastic approach (Arnold & Williams, 1989).

The weather data of the two weather stations were corrected for topography. Potential evapotranspiration (ETP) classes were built based on the effect of elevation, slope and aspect on temperature and radiation and thus, indirectly, on ETP.

Additional data used in this application are the digital soil map 1:50 000 and a digital elevation model (DEM) with a spatial resolution of 31.5 m. The DEM shows significant errors at the edges of the digitized tiles. Thus runoff may appear to leave a subcatchment in the wrong direction.

Land use information was taken from a Landsat image recorded in April 1993 which was classified into six land use classes (urban areas, coniferous forest, deciduous forest, lakes, grassland, agricultural areas). Single pixels and other small areas were eliminated by applying an aggregation algorithm. The number of homogeneous areas was thus reduced significantly. Agricultural statistics at the county level were used to specify the land use of agricultural areas. For each precipitation region a separate annual distribution was assumed. For these land use data no spatial allocation was available, only percentile data.

RESULTS

The regionalization techniques, developed for the test catchments, were then applied to the target catchment of the special research programme, the "obere Leine" catchment (1000 km^2). During the application some new aspects emerged. In comparison to the lowlands where the concept was developed, the lateral flow components became much more important. The limitations of the database became very important necessitating new ideas.

The simulation period was nine years, January 1981–November 1989. Using the regionalization concept, water fluxes of the upper Leine catchment were calculated using a limited number of representative ecotopes. The number of simulation runs needed for the simulations performed are shown in Table 1.

Table 1 Number of representative ecotopes used and simulation runs needed for different hierarchical steps in the cluster analysis. The total number of catchment ecotopes is 9266, requiring 16 631 simulation runs.

Number of appropriate clusters	2	5	8	11	14	18	21	24	30
Number of simulations performed	107	255	401	509	608	751	855	947	1102
% of simulations	0.6	1.5	2.4	3.1	3.7	4.5	5.1	5.7	6.6

An ecotope may cover different ETP- (potential evapotranspiration) or climate-regions. Therefore a greater number of simulation runs were needed than the number of ecotopes. Nevertheless, there was a considerable reduction of computation time. The simulations performed using between two and 30 representative ecotopes needed 0.6 to 6.6% of the simulation runs (16 631) for all the ecotopes (9266). Despite the significant reduction of computation time the spatial resolution of simulation results is preserved. Using 18 representative, ecotopes a stable simulation of the catchment

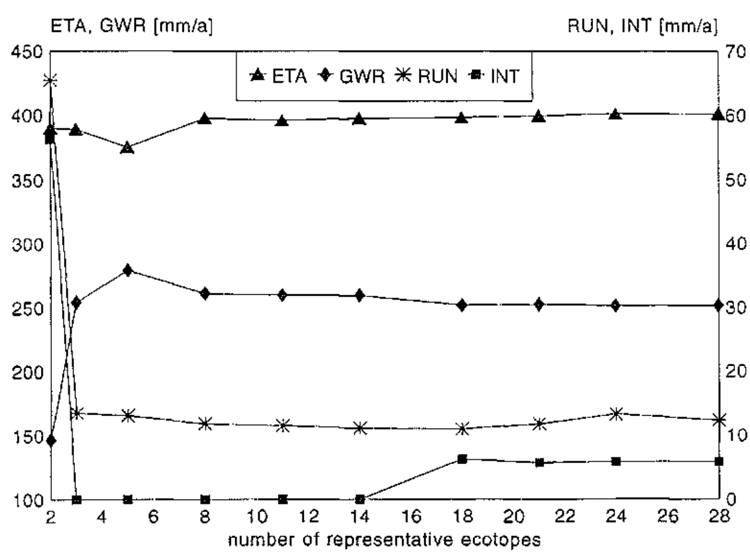

Fig. 2 Development of the calculated catchment water balance depending on the number of representative ecotopes used; ETA = actual evapotranspiration, GWR = groundwater recharge, RUN = surface runoff, INT = interflow.

water fluxes is obtained (Fig. 2). The model system generally underestimates runoff by about 20%. Possible reasons are incorrect precipitation data and the fact that impermeable surfaces are not considered by the model. The trend of underestimation of runoff values becomes conspicuous when comparing measured and simulated monthly discharge values. The correlation is significant (r = 0.89). The trend is reproduced well but measured discharges are generally higher than simulated (Fig. 3).

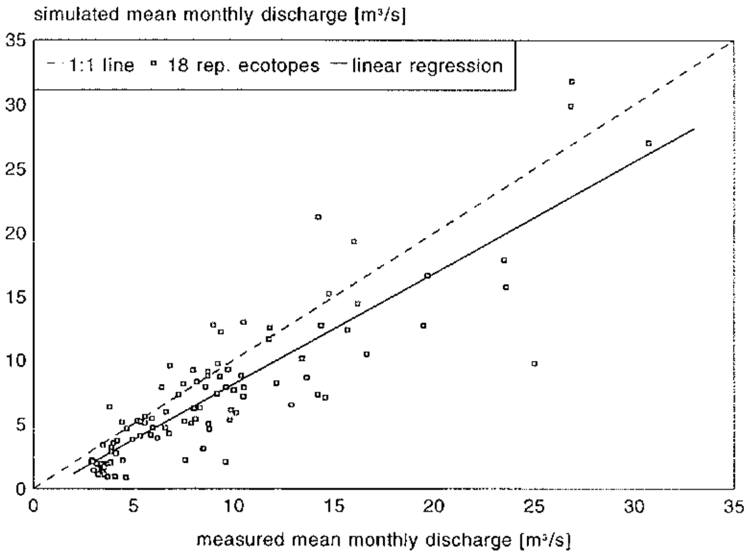

Fig. 3 Linear regression between measured and simulated monthly discharges at the catchment outlet (r = 0.89).

For each time step of the simulation and each site in the catchment, the water fluxes can be specified. Figure 4 shows, for example, the spatial pattern of actual evapotranspiration for the whole drainage area in 1987. Land use is the dominant factor affecting the coarse pattern of dark (lakes, forest) and bright (grassland, agriculture) colours. A dark colour induces high and a bright colour low evapotranspiration values. Further differentiation is caused by soil properties, topography and weather data.

CONCLUSIONS

A regionalization method has been developed which enables the simulation of water fluxes for large drainage areas without having to change the model concept used. Using representative ecotopes, the computer time required is considerably reduced without losing the spatial resolution of the landscape. The results show that the applicability of the concept is limited only by data availability and not by the model itself or the regionalization concept. As the size of an investigated catchment increases the information density decreases. The errors caused by interpolation and disaggregation of data become more and more crucial. Beyond a critical information

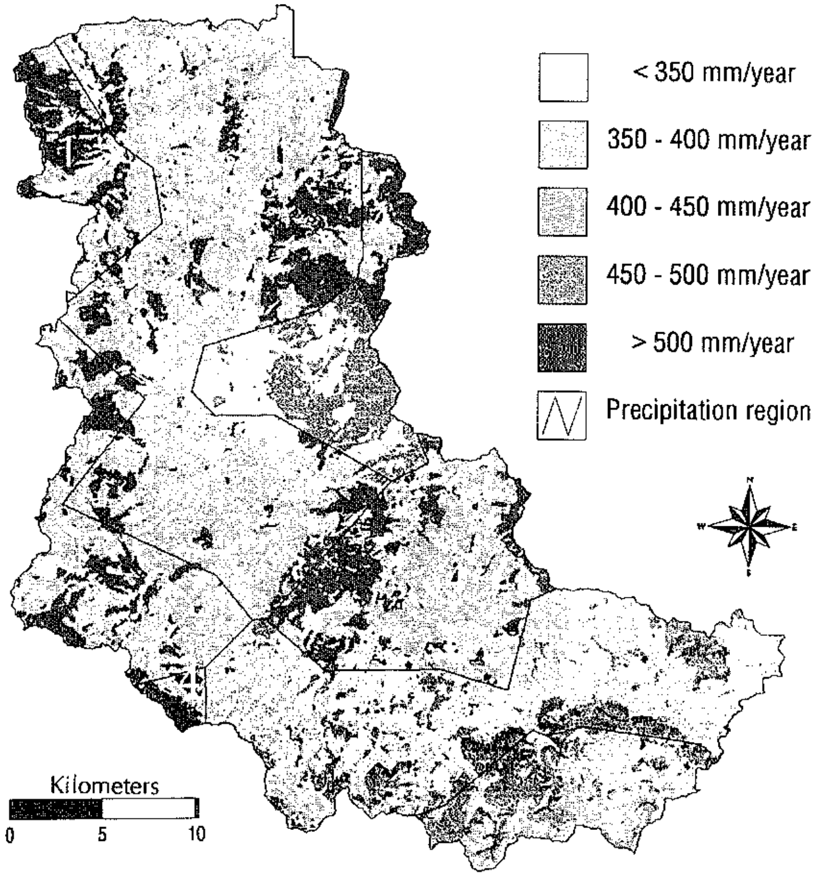

Fig. 4 Areal actual evapotranspiration in 1987 in the target catchment "obere Leine".

density, dynamic modelling using physically based approaches, is no longer advisable. It is difficult to determine the amount of data necessary, but detailed information of land use (e.g. agricultural statistics) and a digital soil map (1:50 000) should be available. Depending on the region, a highly resolved DEM should also be used. For lowland areas, a DEM with a vertical resolution and accuracy of centimetres may be necessary, but not if only regional evapotranspiration is required. In principle the regionalization concept presented is also applicable to mass fluxes, but the regionalization indices would then have to be extended.

Acknowledgement This work was supported by a special research programme of the Deutsche Forschungsgemeinschaft "Regionalization in Hydrology" (DFG-Schwerpunktprogramm "Regionalisierung in der Hydrologie").

REFERENCES

Arnold, J. G. & Williams, J. R. (1989) Stochastic generation of internal storm structure at a point. *Trans. Am. Soc. Agric. Engrs* **32**(1), 161–167.

Blöschl, G. (1996) Scale and scaling in hydrology. Habilitationsschrift (postdoctoral thesis). *Wiener Mitteilungen 132* (ed. by D. Gutknecht). Technical University Vienna, Austria.

Bormann, H., Diekkrüger, B. & Richter, O. (1996) Effects of data availability on estimation of evapotranspiration. *Phys. Chem. Earth* **21**, 171–175.

Bormann, H., Diekkrüger, B. & Richter, O. (1999) Effects of spatial data resolution on the calculation of regional water balances. In: *Regionalization in Hydrology* (ed. by B. Diekkrüger, M. J. Kirkby & U. Schröder) (Proc. Braunschweig Conf., March 1997). IAHS Publ. no. 254 (this volume).

Diekkrüger, B. (1996) SIMULAT – Ein Modellsystem zur Berechnung der Wasser- und Stoffdynamik landwirtschaftlich genutzter Standorte (SIMULAT – a model system for the calculation of water and matter dynamics on agricultural sites, in German). In: *Wasser- und Stoffdynamik in Agrarökosystemen, Sonderforschungsbereich 179* (ed. by O. Richter, D. Söndgerath & B. Diekkrüger), 30–47. Landschaftsökologie und Umweltforschung no. 24.

Diekkrüger, B., Bormann, H., Renschler, C. & Richter, O. (1998) Kurzbericht AG Diekkrüger/Richter (Report of the Diekkrüger/Richter Working Group, in German). In: *Regionalisierung in der Hydrologie: Abschlussbericht des Schwerpunktprogramms 1992-1998* (ed. by A. Becker, B. Diekkrüger, H.-B. Kleeberg, W. Mauser, G. Peschke, R. Schwarze, G. A. Schhultz & U. Streit). VCH/Wiley (in press).

Finke, P. (1992) Spatial variability of soil structure and its impact on transport processes and some associated land qualities. Dissertation. Wageningen, The Netherlands.

Lücke, A. (1997) Regionalisierung des Oberflächenabflusses unter Einsatz Geographischer Informationssysteme und multivariater statistischer Verfahren (Regionalization of surface runoff using GIS and multivariate statistics, in German). Dissertation. Naturwiss. Fakultät der Technical Universitat, Braunschweig, Germany.

Nieschulz, K.-P. (1997) Mathematisch-bodenkundliche Regionalisierung des Bodenwasserhaushaltes (Mathematical-pedological regionalization of the soil water balance, in German). *Landschaftsökologie und Umweltforschung no. 27.*

Richter, O., Söndgerath, D. & Diekkrüger, B. (eds) (1996) *Wasser- und Stoffdynamik in Agrarökosystemen, Sonderforschungsbereich 179.* Abschlussbericht in 3 Bänden. Landschaftsökologie und Umweltforschung no. 24.

Smith, R. E, Goodrich, C. & Woolhiser, D. A. (1990) Areal effective infiltration dynamics for runoff and small catchments. In: *Trans. 14th. Int. Congress of Soil Science*, 22–27. ISSS.

Smith, R. E. & Parlange, J.-Y. (1978) A parameter-efficient hydrologic infiltration model. *Wat. Resour. Res.* **14**(3), 533–538.

Thieken, A. H., Lücke, A. & Diekkrüger, B.(1997) Einfluß von Datenqualität und DHM-Analysen auf die Ergebnisse eines Oberflächenabflußmodells. Beiträge zur Geographischen Informationswissenschaft (Effects of data quality and DEM analysis on the results of a surface runoff model, in German). *Salzburger Geographische Materialien* 26, 153–162.

Wood, E. F. (1995) Heterogeneity and scaling land-atmospheric water and energy fluxes in climate systems. In: *Space and Time Scale Variability and Interdependencies in Hydrological Processes* (ed. by R. A. Feddes), 3–19, Cambridge University Press, UK.