Using Monte Carlo simulations to estimate the surface run-off for a distributed recharge model in a catchment area in Scotland.

M. M. Mansour, A. G. Hughes, M. T. Graham & C. Okoronkwo
British Geological Survey, Wallingford, Oxfordshire, UK, OX11 9RQ
majm@bgs.ac.uk

Abstract Recharge modelling has been undertaken using the distributed recharge model ZOODRM to estimate recharge to an alluvial aquifer in the Borders region of Scotland. The alluvium is flat lying and lies on bedrock which outcrops to form the adjacent hills. Recharge is calculated by the ZOODRM model for a water balance for use in a groundwater flow model of the alluvium. ZOODRM estimates the daily run-off as a percentage of the daily rainfall using run-off coefficients and applies a simple approach to route the run-off to rivers based on an aspect map derived from a topographical map. The comparison of simulated river flows to those observed allows the calibration of the recharge model without using the groundwater flow model. Monte Carlo simulations have been undertaken to estimate the run-off coefficient values of the two topographically different zones and to place confidence limits around these values. The run-off coefficients are either specified (Method 1) or related exponentially to the surface gradients and the antecedent soil moisture conditions (Method 2). The advantages and disadvantages of using these methods are discussed. Both methods produced simulated river flows that are in good agreement with those observed but Method 2 provided a slightly better match because run-off coefficient values were dependent on the antecedent soil moisture.
Keywords recharge modelling, surface run-off, calibration, Monte Carlo simulation, object oriented.

INTRODUCTION

The Howden groundwater source comprises boreholes in valley bottom alluvium which is in hydraulic contact with the Ettrick Water. The source comprises four abstraction boreholes situated next to the Ettrick Water, several kilometres upstream from Selkirk in the Borders region of Scotland (Fig. 1). The site is operated by Scottish Water who are investigating an increase in the amount of water they abstract from the site. Scottish Water commissioned BGS to undertake a hydrogeological study of the groundwater system. This was supported by the construction of a water balance which led to the development of a distributed recharge model using the BGS model ZOODRM. Although the recharge model involves many parameters that describe the behaviour of the recharge processes, the values of most of these parameters are obtained from the available literature. The calibration of the recharge model is dependent mainly on the run-off coefficient values which control the amount of surface water generated.

This paper describes two different approaches used in the recharge model ZOODRM to calculate the run-off component of the surface water. It also describes the use of the Monte Carlo simulation as a tool to determine the values of the run-off coefficients, to calibrate the model, and to estimate the uncertainty.

DESCRIPTION OF THE MODEL

ZOODRM (Mansour and Hughes, 2004) is a distributed recharge model developed using the object-oriented technology. It is possible to incorporate local grid refinement in the model to increase the resolution of the recharge calculations over discrete areas. This also makes the recharge model fully compatible with the groundwater flow model ZOOMQ3D which incorporates local grid refinement (Jackson and Spink, 2004). Five recharge calculation methods are implemented in the model. These are the conventional Soil Moisture Deficit (SMD) method (Penman, 1948 and Grindley, 1967), the Environment Agency / FAO method (Hulme et al., 2002), and specific methods for recharge calculation in semi-arid, urban and irrigated areas. The model also simulates indirect recharge processes such as routing surface water run-off to rivers and ponds and routing water in the unsaturated zone to springs. A run-off coefficient is used to calculate run-off as a percentage of total rainfall at each node. A Digital Terrain Model (DTM) is used to provide slope aspect and route the calculated run-off from the grid node to the surface water feature. The run-off coefficient values can either be held in data files (Method 1) or calculated using an equation that is a function of the topographical gradient and the antecedent conditions of the SMD (Method 2). The mathematical equation used in Method 2 is:

$$\text{Runoff} = \left(1 - e^{\frac{-\text{Gradient Factor}}{100-100*\text{Gradient}}}\right)\left(\frac{\text{SMD}_{\text{Threshold}} - \text{SMD}}{\text{SMD}_{\text{Threshold}}}\right)^2 \quad (1)$$

where Gradient Factor is a constant that controls the relationship between the run-off coefficient value and the gradient (Dimensionless), Gradient is the topographical gradient at a grid node (Dimensionless), $\text{SMD}_{\text{Theshold}}$ is a constant representing a threshold value for the soil moisture deficit, and if it is exceeded no run-off can be generated (mm), and SMD is the soil moisture deficit at a node (mm). In this equation the volume of run-off is positively related to the topographical gradient and inversely related to the SMD.

The recharge model uses a daily time step to undertake the recharge calculations. The topographical characteristics and landuse at a specific location are obtained using calls that extract information from ASCII data files prepared in an ArcGIS environment using the coordinates of that location. Similarly, rainfall and potential evaporation data are obtained at a point from the relevant files using calls that pass location and time.

DESCRIPTION OF THE CALIBRATION METHOD

The model is calibrated by changing the run-off coefficient values using the Monte Carlo simulations (MC). To use MC simulations a mathematical equation is required to vary the run-off coefficient values seasonally when they are defined in the data files (Method 1). The selected equation, used to describe the seasonal variation of the run-off coefficient values, is as follows:

$$\text{Runoff}_{\text{imonth}} = 1 - e^{\frac{a}{\text{imonth}}} \quad (2)$$

where (a) is a constant which controls the shape of the exponential equation (dimensionless), imonth is the month number in the year (from 1 to 6). It is assumed that the seasonal variation of the run-off coefficient values is symmetrical with respect to the first and second halves of the year. This equation is, therefore, applied for the first six months of the year and the same calculated values are repeated but in reverse order for the last six months of the year.

Monte Carlo simulations are applied to generate the constant (a) for Method 1 and the two constants Gradient Factor and $SMD_{Threshold}$ when equation (1) is used in Method 2. For each MC simulation a routine which produces a uniform random deviate between 0 and 1.0 (Press et al., 2002) is used to set the values of (a) or Gradient Factor and $SMD_{Threshold}$. These values are fed to the recharge model so that a recharge simulation can be undertaken. The objective function that determines the acceptance or the rejection of these values is the sum of the squared errors between the observed and simulated river flows. The results are analysed with regard to their frequency of occurrence.

DESCRIPTION OF THE MODELLED AREA

The recharge model boundaries are drawn along the topographic highs defining the catchment area of the Ettrick Water. The model area has an elongated shape, with the two long boundaries trending southwest-northeast parallel to the Ettrick Water (Fig. 1). The two limbs in the south-west of the model area terminate at the upstream end of the Ettrick Water (Ettrick Head) and the upstream end of a tributary. The north-east end of the model is located at a gauging station. The extent of the resulting recharge model area was larger than the area of the groundwater flow model to allow the construction of the water balance.

The modelled area is divided into two topographically different zones (Fig.1). Zone 1 is flat and formed mainly by glacial outwash and alluvial deposits. Zone 2 represents the higher ground, which is mainly associated with the bedrock outcrop. Landuse on the hill slopes is dominated by pastoral farming, while mixed farming is the main landuse on the alluvium.

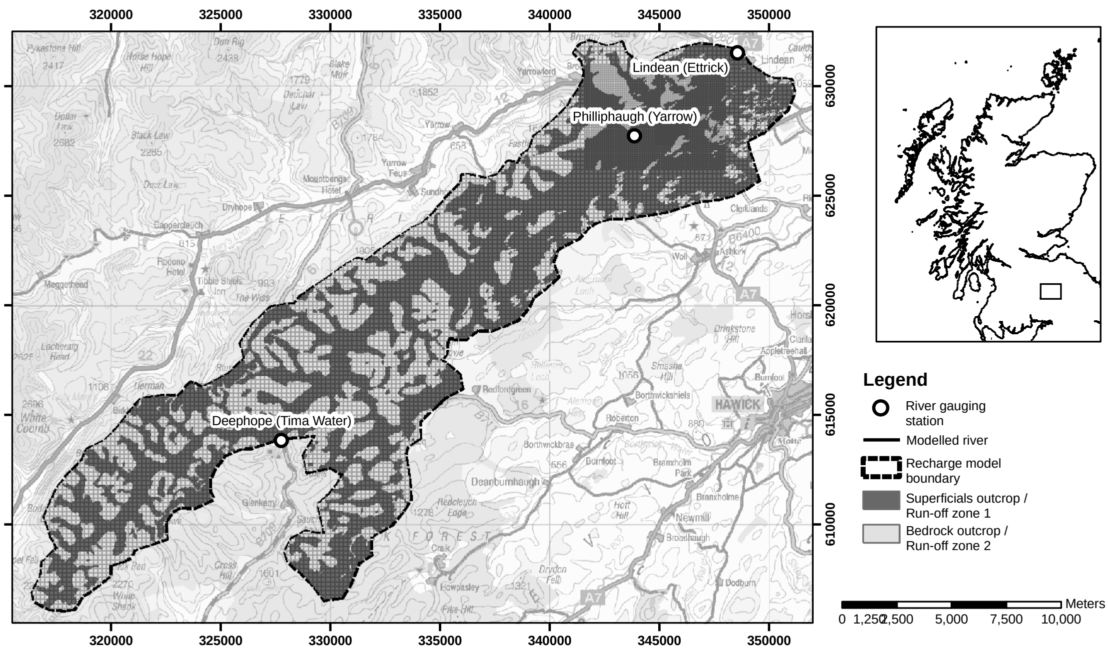

Fig. 1 Part of the Ettrick Water catchment. (This map is based upon Ordnance Survey topographic material with the permission of Ordnance Survey on behalf of The Controller of Her Majesty's Stationery Office, © Crown copyright. Unauthorised reproduction infringes Crown Copyright and may lead to prosecution or civil proceedings. Licence Number: 100017897 [2006])

APPLICATION OF THE MODEL

A model grid with 200 m square cells is used. A refined grid with 100 m square cells is introduced in the ZOOMQ3D numerical flow model to improve the positioning of the wells in the model and the shape of the modelled river near the Howden wellfield area. The same grid refinement is also introduced in the recharge model to calculate the recharge values at the refining grid nodes (Fig. 1).

The CEHDTM map (Morris and Flavin, 1990) is used to derive the topographical gradient values and the aspect directions used in the surface flow routing. Rainfall data at the gauging stations in the area and the MORECS potential evaporation data are obtained from the Centre of Ecology and Hydrology database. Grid nodes are related to the evaporation squares and rainfall stations using Theissen polygons. Landuse is assumed to be the same throughout the study area. This assumption affects the calculated recharge values but has no effect on the surface run-off values. River flows are obtained from the hydrological data book (NERC, 2003). Overland flows are then calculated using the low flow method for baseflow separation (Gustard et al., 1992),. The calibration of the model is undertaken by comparing the simulated and calculated overland flows at the gauging station (Fig. 1).

RESULTS

In total, 5000 Monte Carlo simulations were undertaken to estimate the run-off coefficient values of parameter (a) of equation (2) or the values of the Gradient Factor and the $SMD_{Threshold}$ constants of equation (1). The parameter values of the simulations with objective function values that are less than a user specified value are recorded to a text file.

The values of the two parameters used in equation (2) that produce the least sum of squared errors are 0.1 and 9.6 for the flat (superficial deposits) and hilly (bedrock) run-off zones respectively. These values yield the seasonally varying run-off coefficient values shown in Table 1. Fig. 2 shows a comparison between the calculated overland flows and the simulated ones. This shows that the simulated results are in good agreement with the calculated data although the model failed to produce the high peak flows. In addition, there are simulated flows at times when no overland flows are calculated.

Table 1 Run-off coefficient values calculated using equation (2) and producing the lowest sum of squared errors

Zone	Equation 2 parameter value	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Flat / Glacial deposits	0.1	0.095	0.048	0.033	0.025	0.02	0.017	0.017	0.02	0.025	0.033	0.048	0.095
Hilly / Bedrock outcrop	9.6	0.99	0.99	0.96	0.91	0.85	0.8	0.8	0.85	0.91	0.96	0.99	0.99

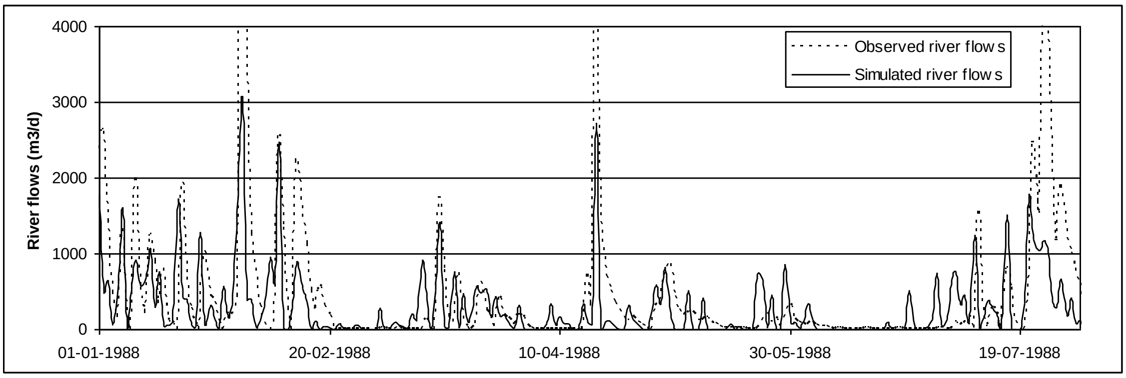

Fig. 2 Calculated versus simulated overland flows. Run-off coefficient values retrieved from files.

The parameter values used in equation (1), which relates the generated run-off water to the SMD and the gradient, and that produce the least sum of the squared errors are shown in Table 2. Fig. 3 shows the comparison between the calculated overland flows and the simulated ones produced using these values. This method of calculating the run-off values does not significantly improve the match between the simulated and calculated peak flows but the match for low river flows is better.

Table 2 Parameter values used in equation (3) to calculate the run-off coefficient values. These parameter values produce the lowest sum of square errors

Zone	Topographical gradient factor	$SMD_{Threshold}$
Flat / Glacial deposits	55.4	34.6
Hilly / Bedrock outcrop	99.7	28.2

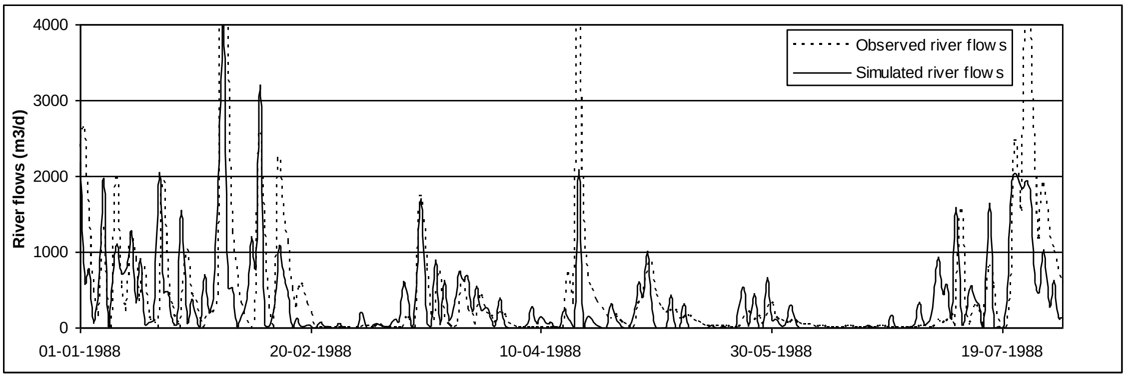

Fig 3 Calculated versus simulated overland flows. Run-off coefficient values related to topographical gradient and SMD values.

Uncertainty associated with these parameter values is examined here by grouping the MC simulations that produce close parameter values into sets and studying the frequency of occurrence of these sets. For Method 1, 300 MC simulations with the least sums of squared errors are examined. These are grouped in 55 sets that have the frequencies of occurrences shown in the histogram in Fig. 4. This figure does not show a normal distribution shape and shows that many sets have the same frequency of occurrence. This indicates that a high level of uncertainty is associated with the estimated parameter values.

For Method 2, 500 MC simulations with the least sums of squared errors are examined. Because there are more parameters in this method (four parameters) than there are in Method 1, these sets are grouped in 420 groups. The frequencies of occurrences of these groups are shown in the histogram in Fig. 5. This indicates that uncertainty is high and that there may be a need to restrict the degree of freedom of the parameter values in the MC simulations.

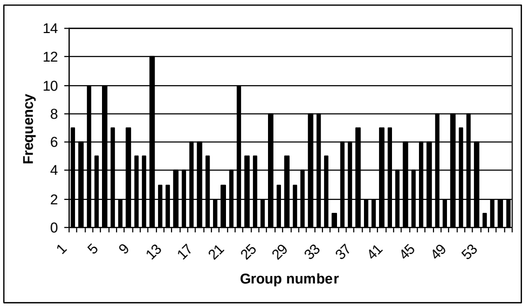

Fig 4 Frequency of occurrences of the sets holding MC simulation results using Method 1

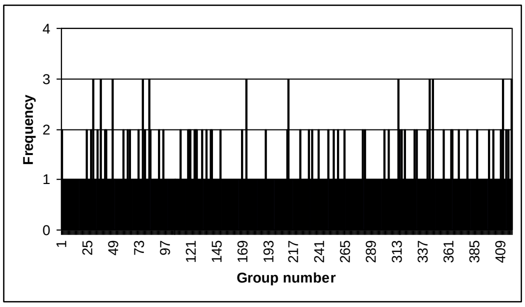

Fig 5 Frequency of occurrences of the sets holding MC simulation results using Method 2

CONCLUSIONS

Two different methods to calculate the volume of surface run-off are implemented in the distributed recharge model ZOODRM. The first method uses run-off coefficient values defined in data files and the second method uses mathematical equations that relate the amount of generated run-off at a location to the topographical gradient and SMD value at that location. Monte Carlo simulations are used to estimate the values of the run-off coefficients or the parameter involved in the mathematical equations. The two methods produce simulated overland flows that are in close agreement with the observed ones, with the second method producing a better match because it offers a more realistic representation of the surface run-off process. The uncertainty analysis, however, shows a wide variation in the parameter values and indicates that a high uncertainty level is associated with the values of these parameters. Limiting the variations of the parameter values in the MC simulations may generate parameter values with higher confidence levels.

Acknowledgements Published with the permission of the Executive Director of the BGS (NERC).

REFERENCES

Grindley J. (1967) The estimation of soil moisture deficits. *Meterol. Mag., **96** (1137), pp 97-108.*

Gustard A., Bullock A. & Dixon J. M. (1992) Low flow estimation in the United Kingdom. *Institute of Hydrlogy Report No. 108, Wallingford, U. K.*

Hulme P., Grout M., Seymour K., Rushton K., Brown L. & Low R. (2002) Groundwater resources modelling: guidance notes and template project brief (Version 1). *Environment Agency, Bristol, U.K.*

Jackson C. R. & Spink A.E.F. (2004) User’s manual for the groundwater flow model ZOOMQ3D. *British Geological Survey Internal Report no. CR/04/140N, Wallingford, U.K.*

Mansour M. M. & Hughes A. G. (2004) User’s manual for the recharge model ZOODRM. *British Geological Survey Internal Report no. CR/04/151N , Wallingford, U.K.*

Morris D. G. & Flavin R. W. (1990) A digital terrain model for hydrology. *Proc 4th International Symposium on Spatial Data Handling. Vol 1 Jul 23-27 Zurich, pp 250-262.*

Natural Environment Research Council (NERC) (2003) Hydrological data United Kingdom, Hydrometric register and statistics 1996-2000. *Centre for Ecology and Hydrology, Wallingford, U.K.*

Penman H.L (1948) Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. London, Ser. A, **193**, pp 120-145.*

Press W. H., Teukolsky S. A., Vetterling W. T. & Flannery B. P. (2002) Numerical recipes in C++ : the art of scientific computing (2nd ed.). *Cambridge University Press, Cambridge, U.K.*