

Simple Overland Flow Calculation Method for Distributed Recharge Models

Mansour MM, Hughes AG, and Barkwith A

Abstract

A method to improve the calculation of overland flow in distributed recharge models is presented and applied to two sub-catchments in the Thames Basin, UK. Recharge calculation studies tend to simulate the run-off flow component of river flow in a simplistic way, often as a fraction of rainfall over a particular period. The method outlined in this study is intended to improve the calculation of recharge estimates in distributed recharge models and is not presented as an alternative to complex overland flow simulators. This method uses seasonally varying coefficients to calculate run-off for specified hydrogeological classes or run-off zones previously used to model base flow index (BFI) variations across the basin, and employs a transfer function model to represent catchment storage. Monte Carlo simulation was applied to refine the run-off values. Decoupling the run-off zones between the two sub-catchments produces a better match between the simulated and observed values; however, the difference between observed run-off and simulated output indicates other factors, such as landuse and topographical characteristics, affecting the generation of run-off flow need to be taken into account when classifying run-off zones.

Introduction

Often, streamflow on river networks with no gauging equipment needs to be estimated without the use of observational data. Previously streamflow hydrographs were used for regional analysis (Hall, 1968 and Tallaksen, 1995), however, more recently the use of numerical modelling techniques has been employed to simulate the contribution of baseflow and surface run-off to fluctuations of streamflow. In numerical distributed recharge calculation models, calculating river flows is a necessary task required to account for water infiltrating into the ground.

On its journey to a river network, a percentage of catchment precipitation is lost to surface storage (interception and depression), infiltration, and evapotranspiration. Losses at this stage are affected by local climate, vegetation type and density, soil type and degree of disturbance, catchment slope, and water course efficiency (MCA, 1997). The remainder completes its journey through a variety of spatially and temporally diverse pathways, as direct surface run-off, interflow, channel precipitation, or groundwater flow. The former three pathways are traditionally combined into *surface flow*, which provides relatively fast transport from deposition into surface watersheds, while groundwater transport operates on longer temporal scales. Groundwater recharge describes the transition of surface water into the saturated zone, while baseflow can be described as the contribution of groundwater into a particular watercourse (stricter definitions are provided by: Hall, 1968; Chow et al., 1988; Ward and Robinson 1990; and Eckhardt, 2008), and is independent of surface flow processes. Baseflow index (BFI) is the temporally averaged ratio of total stream flow to baseflow and therefore represents delayed contribution of groundwater to

streamflow. BFI may be influenced significantly by catchment geology, however, until recently this relationship had not been quantified in a systematic manner (Bloomfield, 2009).

Previous recharge calculation studies have tended to simulate river flow fluctuation (with respect to rainfall) without taking into account the temporal differences of flow pathways. These studies treat the run-off coefficient in a simplistic manner (MCA, 1997) as either; constant fraction, a fraction of rainfall over a particular period (Bogena et al., 2005); constant loss, related to the infiltration capacity of the catchment soil (Rutledge, 1997); or as a combination of the two (initial loss). ZOODRM is a British Geological Survey (BGS) distributed recharge calculation model (Mansour et al., 2004) that calculates run-off as a fraction of rainfall. Studies with increased hydrological and geological complexity usually separate surface flow from baseflow, the latter of which is affected by topography, geology, vegetation, and climatic factors. These factors make the estimation of baseflow difficult, with variables exerting catchment (and often sub-catchment) specific influences. Carlston (1965) and Zecharias and Brutsaert (1988) suggest baseflow is affected by drainage density and average basin slope. Vogel and Kroll (1990 and 1992) agree, and in addition to average basin slope suggested relief was a major influence on baseflow in their study catchment. Nathan and McMahon (1992) studied the effect of variables on baseflow, concluding that in addition to topological influences (area, elevation, and main-stream length), potential evapotranspiration, forest cover, and geology are integral to obtaining an accurate estimate. Lacey and Grayson (1998) suggest geology affects streamflow in at least two ways; baseflow through the storage of groundwater, and surface flow through the formation of soils with differing depth, density, hydraulic properties, and capacity for vegetation. Climate change at the global scale may influence recharge at catchment scale, however, climatic variations at smaller scales (spatial and temporal) affect all aspects of stream flow, creating changes in recharge baseflow, and surface flow (Wilkinson and Cooper, 1993; and Maxwell and Kollet 2008). To account for local seasonal climatic fluctuations, distributed recharge values are needed for the accurate simulation of groundwater flow (Jakeman and Hornberger, 1993; Wilkinson and Cooper, 1993; Cooper et al., 1995; and Erskine and Papaioannou, 1997). Although there are several methods for estimating recharge, a well calibrated model is required to produce accurate estimates of recharge values (Fazal et al., 2004; Ye et al., 2005; and Uddameri and Kuchanur, 2006).

The recharge model used in this study is calibrated by matching simulated stream flow to observational data. While recharge models depend on relatively large number of parameters, the values of which can be set from literature values (see for example: Nash, 1960; Hall 1968; and Tallaksen, 1995), the BFI (and therefore run-off) cannot be measured accurately (Eckhardt, 2008). The run-off coefficients determine the generated surface flow component of total river flow recorded at a gauging station using a simple relationship (see for example: Sandstrom 1995). Studies examining the geological control on BFI indices, often using regression analysis (Leeper and Uren, 1993; Nathan et al., 1996; Lacey and Grayson 1998; Bloomfield et al., 2009), help to define the run-off zones and the range of run-off values used for numerical simulation.

While previous work focuses on simulating whole river flow, the BGS approach is to simulate the run-off flows as accurately as possible. The groundwater flow model will generate the baseflow component of total river flows. A key success of overland flow simulation is the inclusion of catchment (surface) storage especially if the catchment area is relatively large and leads to a time of concentration (the longest time required for a particle of water to travel from the sub-catchment upstream end to the sub-catchment outlet or the location of the gauging station) that is greater than the model time step (Hessel, 1994; Atkinson et al., 2002; and Jothityangkoon et al., 2001). Cross-correlation of rainfall and surface flow values indicates the number of rainfall days that have a direct effect on the gauging station observational data. However, using a transfer function model to relate run-off to rainfall directly may not work because the run-off coefficients involved in such model types fluctuate temporally. Complex transfer function models may describe the relationship between run-off and rainfall after stripping all seasonality behaviour, however, the obtained model may not describe the run-off rainfall relationship explicitly. A process that describes the seasonal variations of the run-off coefficients needs to be included before attempting to fit a transfer model to the simulated data. The IHACRES model (Jakeman et al., 1990 and Jakeman and Hornberger, 1993) , for example, is a hybrid conceptual-metric model and contains a non-linear loss module that converts rainfall into effective rainfall and a linear module that transfers effective rainfall to stream discharge. The non-linear module precedes the linear routing module to calculate the total stream flow and includes parameters related to the soil moisture index and drying rate.

The aim of this study is to investigate and implement a simple approach to calculating accurate surface flow values in a distributed numerical recharge model. The study uses the BGS ZOODRM model and is focused on river flows recorded at two gauging stations in the Thames Basin, UK. It benefits, therefore, from the hydrogeologically-based litho-stratigraphic classification zones suggested by Bloomfield et al. (2009) previously used to model BFI variations across the basin. A complimentary aim is to investigate the validity of BFI values obtained from regression analysis (Bloomfield et al., 2009), which relates lithological catchment characteristics and BFIs in observed run-off flows values.

Description of the recharge model

BGS developed the distributed recharge model ZOODRM to produce distributed recharge values over grids that are compatible with those of the BGS flow model ZOOMQ3D. The latter model incorporates local grid refinement in a Cartesian mesh to improve the simulation of groundwater flows at smaller scales (Jackson and Spink, 2004). ZOODRM also uses the ZOOMQ3D pre-processor application ZETUP (Jackson, 2004) to create numerical grids and transform river spline points into numerical rivers. There are two levels of numerical grids in the recharge model. The upper level grid is used to represent topographical characteristics of the catchment and surface flow processes. The lower level grid is used to simulate the movement of percolated water in the soil zone, which can be vertical or lateral. Soil zone water moving laterally discharges into water features such as rivers or springs. Seepage of soil zone water along lower elevation ground surfaces is not permitted in the model. The lateral movement is based on user specified hydraulic gradients and groundwater velocities, which can be obtained from separate groundwater models. A

time delay can be specified at the lower grid nodes to control the rate that percolated water leaves the soil zone and becomes part of the water table. The purpose of this simplified representation of water movement in the soil zone is to improve the calculation of the water balance.

Direct and indirect recharge can be accounted for in the recharge model. Direct recharge calculations are performed using either the soil moisture deficit method (SMD), based on the work of Penman (1948) and Grindley (1967), or the Environment Agency-FAO method (Hulme et al., 2002). The SMD method calculates potential recharge based on rainfall and potential evaporation values and the characteristics of land-use represented by vegetation wilting point and root constant. The EA-FAO method employs a more sophisticated approach to calculate the recharge values. This approach involves soil type, soil wetness, and a better representation of seasonal plant growth than that used in the SMD method. Additional methods for the calculation of recharge in urban areas and semi-arid areas are also developed. The recharge calculation method in semi-arid areas is based on the work of Lange et al. (2003) and is used in a study in the West Bank, Palestine (Hughes et al. 2008).

ZOODRM can use rainfall and evaporation data recorded at weather stations to distribute the rainfall and evaporation values over the study area. The distribution of these values is based on the use of Thiessen polygons and maps of average rainfall and evaporation values. An alternative method is to use distributed daily rainfall and evaporation data in gridded ASCII format. The model reads point data from ASCII gridded maps giving rainfall and evaporation as well as geological, topographical, and landuse information. Majority land-use data can be used, or if spatially distributed percentage land-use types are available, they can be processed by the model to assign more than one land-use type at each grid node. Recharge values are then calculated based on these percentages.

ZOODRM is able to represent many of the flow processes that describe the onward transport of rainfall arriving at the ground surface (Hughes et al. 2008) and include: surface flow processes, groundwater flow to shallow spring systems, and a simple method that represents the lateral movement of percolated water in the soil zone based on specified groundwater velocities. The catchment area is split into zones exhibiting different run-off characteristics. A run-off coefficient is allocated to every run-off zone and overland flows are calculated using the run-off coefficient values. Run-off water is routed to downstream nodes using the aspect directions of grid nodes, which are produced from a topographical map in the ArcGIS environment.

Run-off coefficient values can be varied in a simplistic manner by specifying monthly values in dedicated data files. Alternatively, the run-off coefficient values can be related to rainfall intensity and level of soil moisture deficit. A relationship proposed by Rushton (2003) is implemented in the code, however, sophisticated mathematical representations of this relationship are possible and can be used by the model. A study in Howden, Scotland (Mansour et al. 2007) uses a non-linear relationship between rainfall and run-off, and implements Monte Carlo (MC) simulation in an attempt to estimate the value of parameters describing this relationship. However, it has been found that this complicated representation did not greatly improve the calculation of overland flow and the use of the

simple method, which varied the coefficient values seasonally, produced the same quality of results. The main problem faced in this case is that the simulated overland flow exhibits a flashy response and only reflects the rainfall pattern to a certain degree, whilst the observed stream flows possess smooth behaviour as a result of the gradual discharge of surface flow to streams. In both cases the catchment storage was not represented, which made the comparison between the two methods difficult.

Methodology

The run-off flow calculation method presented here consists of two steps. In the first step overland flow is calculated by proportioning daily rainfall using run-off coefficient values specified in data files. Each data file is related to one run-off zone and holds twelve run-off coefficient values, one for each month of the year. This approach simplifies the representation of complex processes used to generate overland flow and which are related to many climate factors, such as soil conditions, topographical characteristics of the area, landuse. At a given point within the recharge model, the calculated overland flow time series reflects the behaviour of the rainfall time series exactly. In fact, the run-off time series is the rainfall time series but reduced by the specified run-off coefficient values (Figure 2). The total overland flow calculated at a gauging station may have a more complicated relation to rainfall because it originates from many locations with different rainfall time series and different run-off coefficient values (Figure 2).

In the second step the effects of catchment storage are included by applying a mathematical equation that describes observed run-off flow for a specified day as a fraction of daily simulated run-off flow plus fractions of simulated run-off flows for previous days. The mathematical equation takes the form of a transfer function that correlates simulated overland flows to observed ones. The parameters involved in the transfer function and the multiplication of simulated run-off coefficients are time independent, and must add up to one, therefore maintaining the system water balance. The transfer function can be derived using the univariate autoregressive integrated moving average (ARIMA) model readily available in the R statistical language (The R Project for Statistical computing). ARIMA relates the value of a vector at a time step to its values at precedent time steps and to a noise model. However, the implementation of this model in the R Language allows the inclusion of an optional vector that is used as a regressor to which the model parameters are fitted. In the current application the univariate time series is the observed run-off flow and the regressor vector is the simulated run-off flow.

The order of the transfer function is determined by cross-correlating the rainfall and observed overland flow. Many simulations involving different transfer functions with different orders are performed and the transfer function model yielding best results is selected. The major drawback of this approach, however, is that run-off coefficient values have to be overestimated so that the resulting overland water can be distributed over several days. A direct comparison between the simulated and observed overland flows is not possible before the application of the transfer function model. To overcome this problem, a new time series of the simulated flows is produced using the moving average functionality available in many statistical packages (for example, Microsoft Excel) and

compare these to observed flows. Observed surface flow components are determined from the total recorded river flow at a gauging station by applying the Institute of Hydrology low flow method for baseflow separation (Gustard et al., 1992).

A trial and error approach is firstly used to select run-off coefficient values and a visual inspection used to check the agreement between the time series obtained from the application of the moving average function to the simulated flow and observed values. The order of the moving average equation is the same as that selected for the transfer function. Once an acceptable match is achieved, the ARIMA model is applied to the calculated coefficient values involved in the regression equation. It is anticipated that the sum of these coefficients is equal or close to unity. If this is not the case, the values of these coefficients are altered, based on user judgement, to sum to unity and therefore maintain the water balance. The final step is to undertake MC simulation, to refine the run-off coefficient values, and to undertake uncertainty analysis. While run-off coefficient values are allowed to vary freely or according to the recommendations proposed by other studies (see for example; Bloomfield et al., 2009), the parameter values in the ARIMA model are fixed to those calculated in the preceding step.

Study Area

Bloomfield et al. (2009) use the Thames Basin, UK, to illustrate the application of regression analysis in examination of geological controls on baseflow index (BFI). They mapped lithological classes onto four hydrogeological classes consisting of; low permeability surficial deposits, aquitards, intergranular aquifers, and fractured aquifers and calculate a baseflow index for each class (Table 1). The four class hydrogeological map produced by Bloomfield et al. (2009) can be used to divide the area into zones with different run-off flow characteristics. Over 150 gauging stations are available on the River Thames (Natural Environmental Research Centre, 2003); however, Bloomfield et al. (2009) screen the number down to 44 gauging stations based on the quality of records and significant anthropogenic influences. Of these 44 gauging stations only 2 are selected for the purpose of this study. Both are located to the north west of the Thames Basin, namely the Enslow Mill gauging station located on the River Cherwell with a grid reference of SP 482183 and the Cassington Mill gauging station located on the River Evenlode with a grid reference of SP 448099. These gauging stations are selected because they have different degrees of complexity associated with their run-off zones. The majority of the catchment area related to the first station, Enslow Mill, is covered by the non-aquifer run-off zone (see Figure 1). The Cassington Mill catchment area is covered by the four classes defined by Bloomfield et al. (2009). The difference in run-off zoning allows the success or failure of the proposed method with respect to the number of involved run-off zones to be discussed.

The boundaries of the two catchment areas are defined based on the topographical highs obtained from a digital terrain model (DTM) (Morris and Flavin, 1990). The catchment area related to Enslow Mill is approximately 547 Km² and the catchment related to Cassington Mill is approximately 441 km². Daily rainfall data at the 20 rainfall stations covering the study area are obtained from the Centre of Ecology and Hydrology, UK. The mean annual

rainfall is approximately 685 mm a^{-1} . Monthly potential evaporation values are obtained from the Meteorological Offices's MORECS data set for squares 137, 148, and 149. The mean annual potential evaporation is approximately 610 mm a^{-1} . The mean flows at Enslow Mill and Cassington Mill are 332 and 322 Ml d^{-1} respectively with baseflow index of 0.65 at Enslow Mill and baseflow index of 0.71 at Cassington Mill (Natural Environment Research Council, 2003).

Ground elevation varies from approximately 240 m AOD at the upstream end of the catchments down to approximately 65 m AOD at the downstream end of the rivers. The land-use in both catchments is predominantly rural, with the exception of the major towns of Banbury, Chipping Norton and Woodstock. The M40 motorway runs up through the east catchment and there are a number of military and private airfields.

Model Application

Distributed recharge values are calculated on a grid with $1,000 \text{ m}$ square cells. Ten landuse classes obtained from the Land Cover Map LCM2000 (Natural Environmental Research Council, 2000) were used to represent landuse in the model. Gridded maps, giving percentage value of each class, are implemented in the model so one grid node may have more than one associated landuse type. Actual Evapo-transpiration and recharge are calculated as a percentage of landuse type using the soil moisture deficit method, which is based on the work of Penman (1948) and Grindley (1967). The value of the root constant and wilting point parameters, which control the amount of evapotranspiration, were obtained from Lerner et al. (1990).

Surface water generated from rainfall at one grid node is transported to adjacent nodes based on the aspect direction of the node. Aspect directions are derived from the DTM in the ArcGIS environment. However, DTM accuracy may be ineffective where the recharge model scale is inappropriate; consequently surface water may get routed to virtual ponds if a river node is not reached. Manual correction to the aspect direction map is necessary to eliminate these virtual pond nodes.

Bloomfield et al. (2009) give values for the four baseflow indices of the specified geologies. The run-off coefficient values related to the defined run-off zones can be derived from the values in Table 1. These are used as initial run-off coefficient values in the basic simulation and are modified, by trial and error, to improve the match between the simulated and observed overland flows. Monthly variations of these coefficients are produced using the sine function, which sets high run-off coefficient values for December, January, and February and low run-off coefficient values for June and July (see Figure 4 for example). The amplitude of the sine wave is selected during the trial and error process. The outcome of this modelling step is a time series of daily overland flows that exhibit a flashy response dictated by the rainfall time series (See Figure 2).

Four ARIMA models with order of 3, 5, 8, and 10 are applied to correlate the simulated and observed flow, accounting for catchment storage. An ARIMA model with an order of 3 indicates the final simulated overland flow on one day consists of three components. These

components are fractions of the flows simulated over the same day and the previous two days. It has been found that increasing the order of the ARIMA model to 8 improved the fit between observed and simulated flows. An ARIMA model with an order of 10 did not provide any noticeable matching improvement over a model with an order of 8 (see Figure 3) and consequently 8 parameters are used in the ARIMA model for each catchment. This means that run-off observed on a particular day is made of run-off components generated up to one week earlier. Table 2 shows the parameter values used for the ARIMA model applied for both catchments. It is clear that the correlation of the current day overland flow to the flows generated on previous days decreases with increased time. For example, Table 2 indicates that 27% of the overland flow related to Enslow Mill is generated on the same simulation day, 14% of this flow is generated the day before, and 6% is generated seven days earlier.

MC simulation was undertaken to refine run-off coefficient values. The task assigned to MC simulation was to select the initial run-off coefficient value for each of the four zones as well as to select amplitude values for the sine equations describing monthly changes to run-off value. The ARIMA parameter values are fixed to those produced after the trial and error step and shown in Table 2. The objective function was to minimise the sum of the squared errors between the simulated and observed overland flows at both gauging stations. Figure 6 shows a comparison between the observed run-off flow time series and the time series of the simulated flows obtained from two different runs for the years 1981 and 1982. The horizontal axis of Figure 6 represents date and the vertical axis represents overland flow in MI day^{-1} . The simulated first run results are derived using the run-off coefficient values, shown in Figure 4, that produce the smallest error in the MC simulations. The simulated results of the second run are derived using the run-off coefficient values used in the final trial and error run (Figure 5).

The simulated time series curves shown in Figure 6 possess smoother behaviour than that of the simulated time series shown in Figure 2. This behaviour is closer to the observed time series behaviour due to inclusion of catchment storage. This can be clearly expressed by the reduction of the root mean squared error (RMSE) between the observed runoff flows and the simulated ones which is 335 MI day^{-1} for the results shown in Figure 2 and 185 MI day^{-1} for the results shown in Figure 6 calculated over the year 1981. When average observed flow values (117 MI day^{-1} for Enslow Mill and 92 MI day^{-1} for Cassington Mill) are compared to the average flow values produced by MC simulation (167 MI day^{-1} for Enslow Mill and 63 MI day^{-1} for Cassington Mill), it is clear that the match between the MC simulated results and the observed ones is better at the Enslow Mill than at Cassington Mill. The difference in simulation response arises, in part, from differences in the number of catchment run-off values. The MC simulation approach creates a compromise solution that produces the best combined results for the gauging stations. The Enslow Mill catchment is composed predominately from a single run-off value (Zone 1), while the catchment that feeds Cassington Mill consists of four run-off values. Trial and error runs produced, on the other hand, produced better match between the simulated and observed results at Cassington Mill (101 MI day^{-1}) as shown in Figure 6, but the fit to the observed data at Enslow Mill (202 MI day^{-1}) is worsened in comparison to the MC simulated output. A better fit between observed and simulated run-off flows can be obtained, therefore, if the run-off zones of the

two catchments are decoupled, i.e. the run-off coefficient values of a class at Enslow are different to those specified for the same class at Cassington Mill.

Discussion and Summary

A simple approach to surface flow calculation is proposed and implemented using the numerical distributed recharge model ZOODRM. This approach calculates overland flow as a fraction of rainfall using run-off coefficients and includes catchment storage by applying a transfer function model. The overland flow calculation procedure is applied to two catchments in the Thames Basin, UK. The simulated overland flows are compared to the run-off components of the river flows recorded at the two gauging stations of these catchments. The run-off component of the recorded river flows is obtained using the Institute of Hydrology baseflow separation technique (Gustard et al., 1992).

Initial run-off coefficient values are obtained from the study undertaken by Bloomfield et al. (2009), which relates lithological catchment characteristics and BFIs in observed run-off flows values. These BFI values were obtained based on the analysis of 44 sub-catchments across the Thames Basin. They represent, therefore, average BFI values and are not expected to be entirely consistent with values calculated in this study; however, they maintain the same relationship to the classified geologies.

The modelled catchments share the same geological classes, the latter of which were given the same run-off coefficient values. This study shows that assigning different run-off values for the same run-off class at different catchments is necessary to improve the match between observed and simulated results. This indicates that the run-off classes assumed to be the same in both catchments are in fact different. This highlights deficiency in the surface water generation approach described, to represent the real complex processes of surface flow generation in an individual catchment. Consequently, the approach used here to calculate run-off flow may not benefit fully from the statistical studies that relate the values of run-off coefficient and BFI to geological classes, such that undertaken by Bloomfield et al. (2009) across the basin as a whole. If the classification of run-off zones followed in this study is correct, the results indicate that additional factors affecting the generation of overland flows need to be accounted for in the proposed modelling approach.

The simulated overland flows produced by the proposed approach have better agreement with the observed overland flows than those produced without the implementation of the transfer function model. The latter flows possess a flashy behaviour similar to the behaviour of the rainfall time series. The inclusion of catchment storage in the modelling procedure yield values for the run-off coefficients that are greater than the values produced without any catchment storage. This has a direct impact on the values of the recharge estimates which are calculated from the amount of rainfall after discounting the surface run-off component. The surface run-off calculation procedure presented is not intended to be used as a numerical approach to replace sophisticated methods developed for surface water numerical models, rather the ultimate objective of the study was to better account for overland flow in a numerical model designed to calculate accurate recharge values.

References

- Atkinson, S. E., Woods, R. A., and Sivapalan, M., 2002: Climate and landscape controls on water balance model complexity over changing timescales, *Water Resources Research*, 38 (12), 1314.
- Bloomfield, J. P., Allen, D. J., and Griffiths, K. J., 2009: Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK. *Journal of Hydrology*, 373, 164-176.
- Bogena, H., Kunkel, R., Schobel, T., Schrey, H. B., and Wendland, F., 2005: Distributed modeling of groundwater recharge at the macroscale. *Ecological Modelling*, 187, 15-26.
- Carlston, C. W., 1965: The effect of climate on drainage density and streamflow. *International Association of Science Hydrological Bulletin*, 11 (3), 62-69.
- Chow, V. T., Maidment, D. R., Mays, L. W., 1988: *Applied Hydrology*. McGraw-Hill, New York, USA.
- Cooper, D. M., Wilkinson, W. B., and Arnell, N. W., 1995: The effects of climate changes on aquifer storage and river baseflow. *Hydrological*, 40 (5), 615-631.
- Eckhardt, K., 2008: A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of hydrology*, 352 (1-2), 168-173.
- Erskine, A. D., and Papaioannou, A., 1997: The use of aquifer response rate in the assessment of groundwater resources. *Journal of Hydrology*, 202 (1-4), 373-391.
- Fazal, M. A., Imaizumi, M., Ishida, S., Kawachi, T., and Tsuchihara, T., 2005: Estimating groundwater recharge using the SMAR conceptual model calibrated by genetic algorithm. *Journal of Hydrology*, 303, 56-78.
- Grindley, J., 1967. The estimation of soil moisture deficits. *Meteorol. Mag.*, 96, 1137, 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.
- Hall, F. R., 1968: Baseflow recession – a review. *Water Resources Research*, 4, 973-983.
- Hessel, R., 2004: Effects of grid cell size and time step length on simulation results of the Limburg soil erosion model (LISEM). *Hydrological Processes*, 19 (15), 3037-3049.
- Hughes A.G., Mansour M.M. and Robins N.S., 2008. Evaluation of distributed recharge in an upland semi-arid karst system: the West Bank Mountain Aquifer, *Hydrogeology Journal*. Doi 10.1007/s10040-008-0273-6.

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. and Spink, A.E.F., 2004. User's manual for the groundwater flow model ZOOMQ3D. British Geological Survey Internal Report IR/04/140. Keyworth, Nottingham, UK.: British Geological Survey.

Jackson, C.R. and Spink, A.E.F., 2004. User's manual for ZETUP, the set up program for the groundwater flow model ZOOMQ3D. British Geological Survey Internal Report IR/04/139. Keyworth, Nottingham, UK.: British Geological Survey.

Jakeman, A. J., Hornberger, G. M., 1993: How much complexity is warranted in a rainfall-run-off model? *Water Resources Research*, 29 (8), 2637-2649.

Jothityangkoon, C., Sivapalan, M. and Farmer, D. L., 2001: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development. *Journal of Hydrology*, 254, 174-198.

Lacey, G. C., and Grayson, R. B., 1998: Relating baseflow to catchment properties in southeastern Australia. *Journal of Hydrology*, 204, 231-250.

Lange J, Greenbaum N, Husary S, Ghanem M, Leibundgut C and Schick A P, 2003, Run-off generation from successive simulated rainfalls on a rocky, semi-arid Mediterranean hillslope, *Hydrological Processes*, 17, 279-296.

Leeper, G. W., and Uren, N. C., 1993: *Soil Science*. Melbourne University Press, Australia.

Maxwell, R. M., and Kollet, S. J., 2008: Interdependence of groundwater dynamics and land-energy feedbacks under climate change. *Nature Geoscience*, 1, 665-669.

Lerner D.N., Issar A.S., Simmers I., 1990. *Groundwater recharge: a Guide to understanding and estimating natural recharge*. IAH, 8.

MCA, 1997: *Minesite Water Management Handbook*. Minerals Council of Australia.

Nash, J. E., 1960: A unit hydrograph study with particular reference to British catchments. *Proceedings of ICE*, 17, 249-282.

Nathan, R. J., Austin, K., Crawford, D., and Jayasuriya, N., 1996: The estimation of monthly yield in ungauged catchments using a lumped conceptual model. *Australian Journal of Water Resources*, 1 (2), 65-74.

Mansour M.M. and Hughes A.G. 2004. User's manual for the recharge model ZOODRM. British Geological Survey Internal Report IR/04/150. Keyworth, Nottingham, UK.: British Geological Survey.

Mansour M.M., Hughes A.G., Malcolm G.T. and Okoronkwo C., 2007b. Using Monte Carlo simulations to estimate the surface run-off for a distributed recharge model in a catchment area in Scotland. ModelCARE 2007, Copenhagen, Denmark.

Mansour M.M., Hughes A.G., Ó Dochartaigh B.E., and Graham M.T., 2008. Representation of Urban Recharge Processes in the Distributed Recharge Model (ZOODRM) of the Clyde Gateway, Scotland. MODFLOW and More 2008: Ground Water and Public Policy – Conference Proceedings, Poeter, Hill and Zheng. Colorado School of Mines, USA.

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.

Nathan, R. J., and McMahon, T. A., 1992: Estimating low flow characteristics in ungauged catchments. *Water Resources Management*, 6, 85-100.

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, *Proceedings of the Royal Society London, Series A*, 193, 120-145

The R Project for Statistical Computing. <http://www.r-project.org/>

Rutledge, A. T., 1997: Model-Estimated Ground-Water Recharge and Hydrograph of Ground-Water Discharge to a Stream. USGS Water-Resources Investigations Report 97-4253.

Rushton, K.R., 2003. *Groundwater Hydrology. Conceptual and Computational Models*. John Wiley & Sons Ltd, West Sussex, England.

Sandstrom, K., 1995: Differences in groundwater response to deforestation: a continuum of interactions between hydroclimate, landscape characteristics and time. *Geojournal*, 35 (4), 539-546.

Tallaksen, L. M., 1995: A review of baseflow recession analysis. *Journal of Hydrology*, 165, 349-370.

Uddameri, V. and Kuchanur, M., 2006: Estimating aquifer recharge in Mission River Watershed, Texas - model development and calibration using genetic algorithms. *Environmental Geology*, 51 (6), 897-910.

Vogel, R. M., and Kroll, C. N., 1990: Generalized low-flow frequency relationships for ungauged sites. *Water Resources Bulletin*, 26 (2), 241-253.

Vogel, R. M., and Kroll, C. N., 1992: Regional Geohydrologic-Geomorphic Relationships for the Estimation of Low-Flow Statistics. *Water Resources Research*, 28 (9), 2451-2458.

Ward, R. C., and Robinson, M., 1990: *Principles of Hydrology*. McGraw-Hill, New York, USA.

Wilkinson, W. B., and Cooper, D. M., 1993: The response of idealised aquifer/river systems to climate change. *Hydrological Sciences Journal*, 38 (5), 379-390.

Ye, M., Pohlmann, K., Chapman, J., and Shafer, D., 2005: On evaluation of recharge model uncertainty: a priori and a posteriori. *International High-Level Radioactive Waste Management Conference*, April 30 – May 4, Las Vegas, NV.

Zecharias, Y. B., and Brutsaert, W., 1988: The influence of basin morphology on groundwater outflow. *Water Resources Research*, 24 (10), 1645-1650.

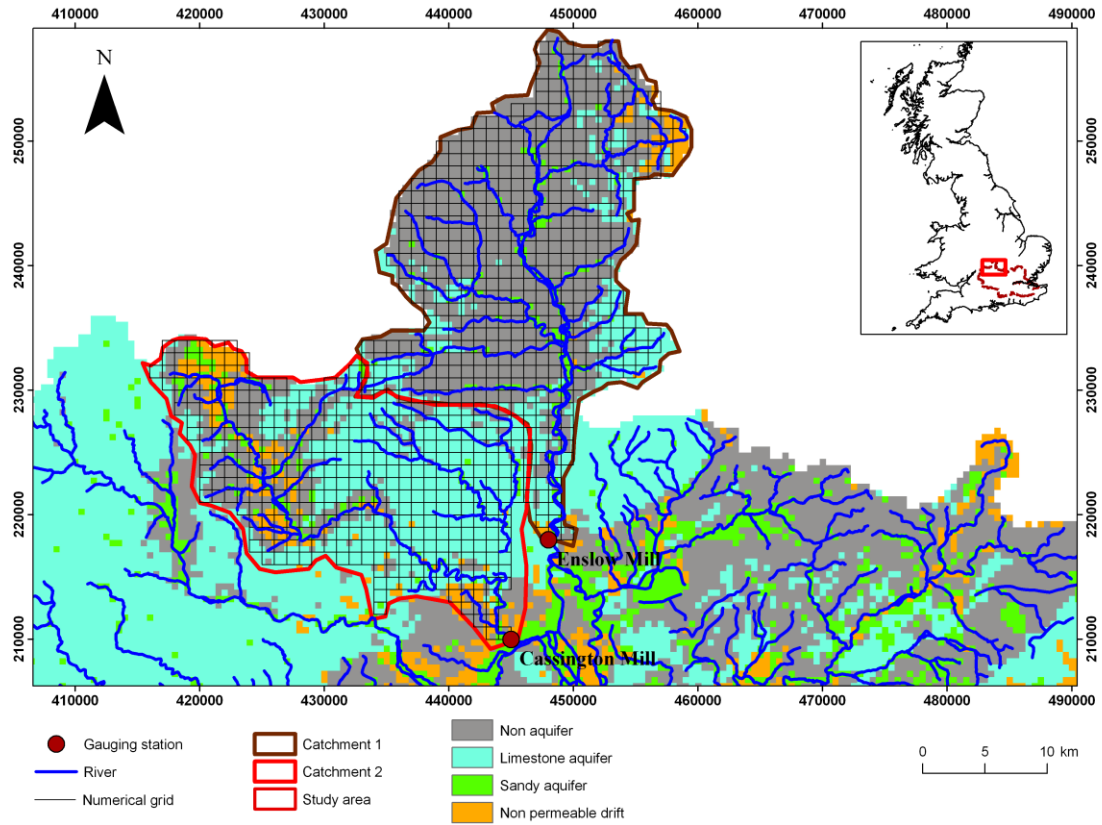


Figure 1 Location of the study area

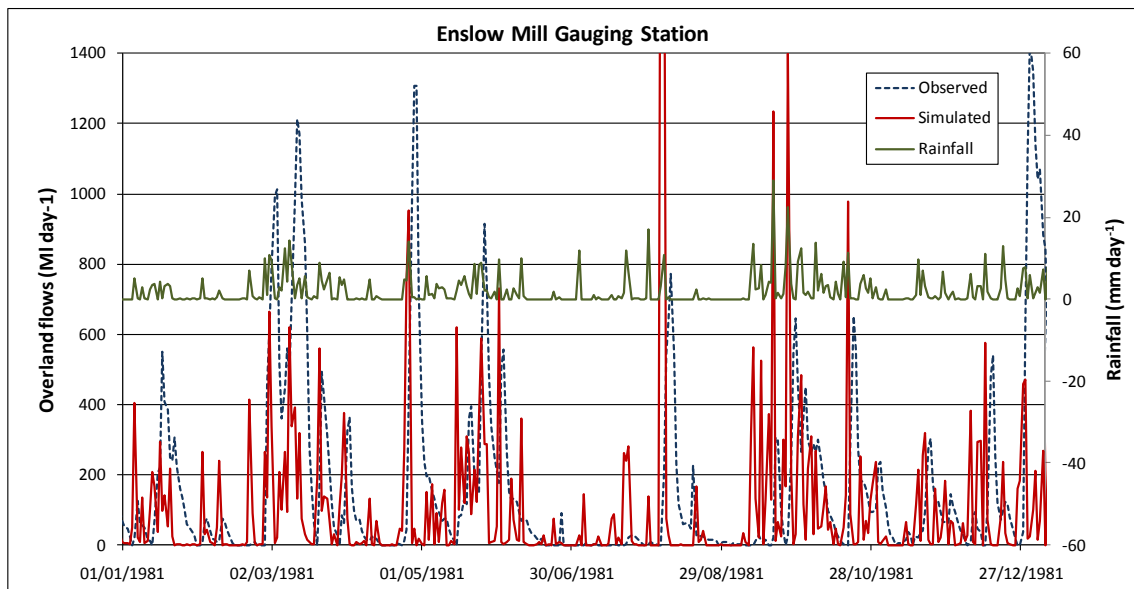


Figure 2 Example of overland flows fitting without the inclusion of catchment storage

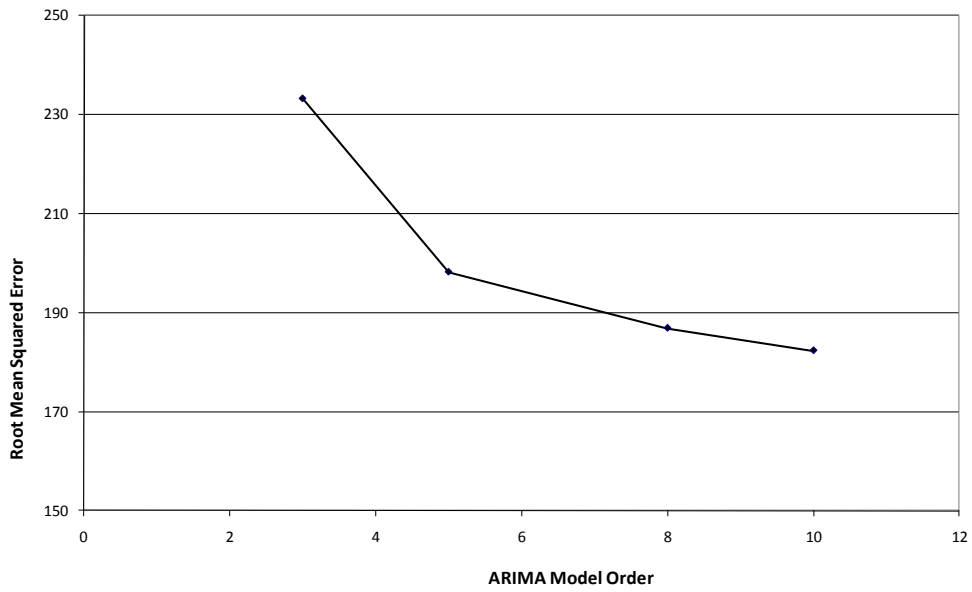


Figure 3 Root mean squared error of results from ARIMA models with different orders

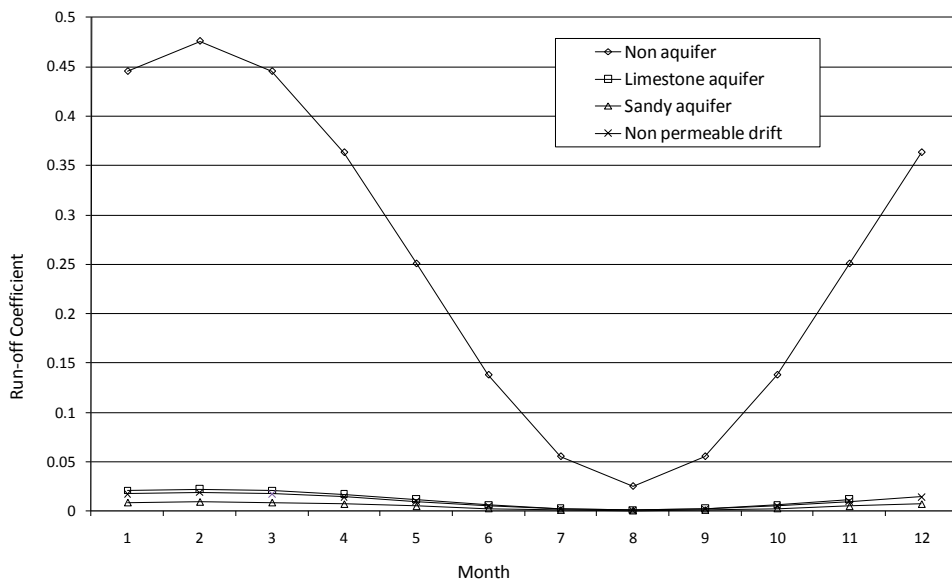


Figure 4 Run-off coefficient values produced from Monte Carlo simulations.

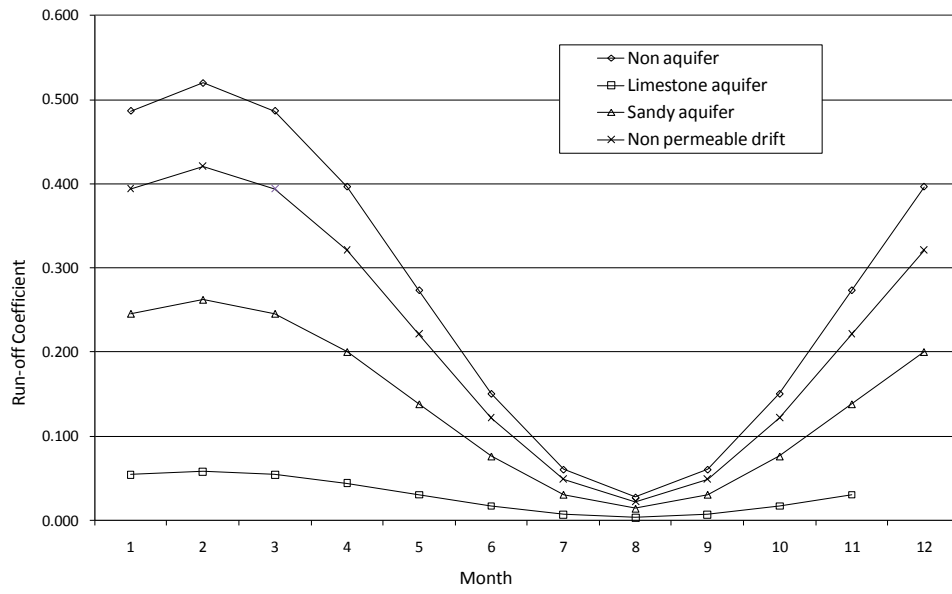


Figure 5 Run-off coefficient values produced from a trial and error run.

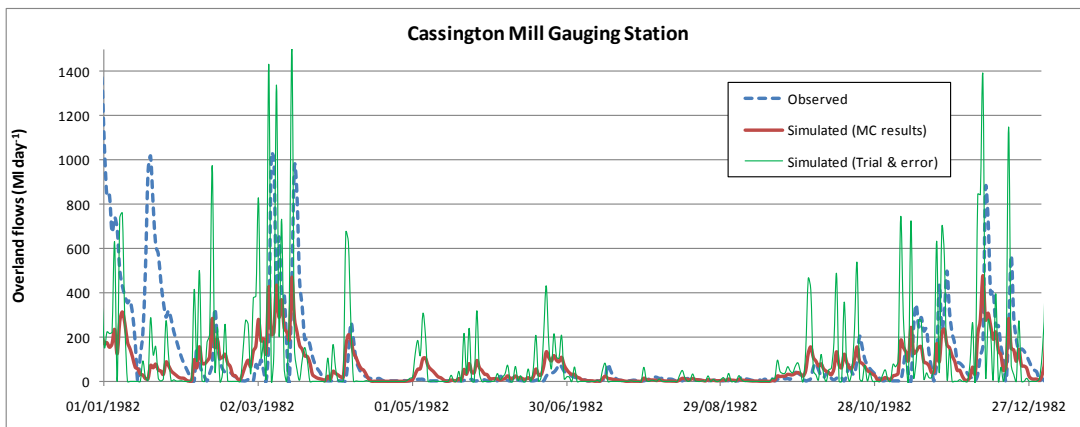
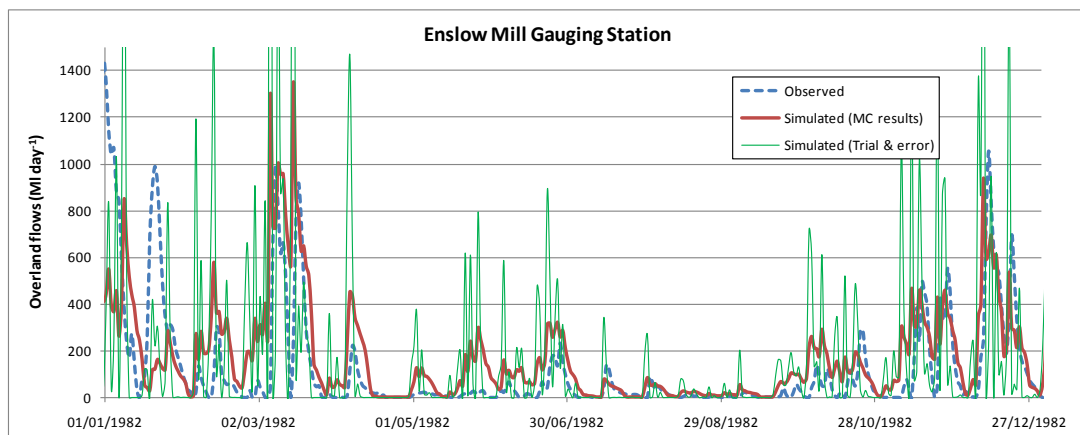
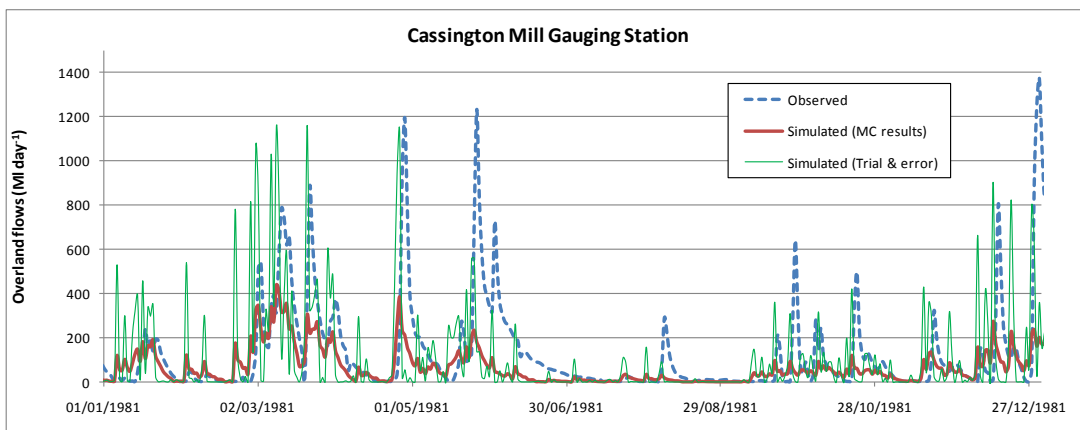
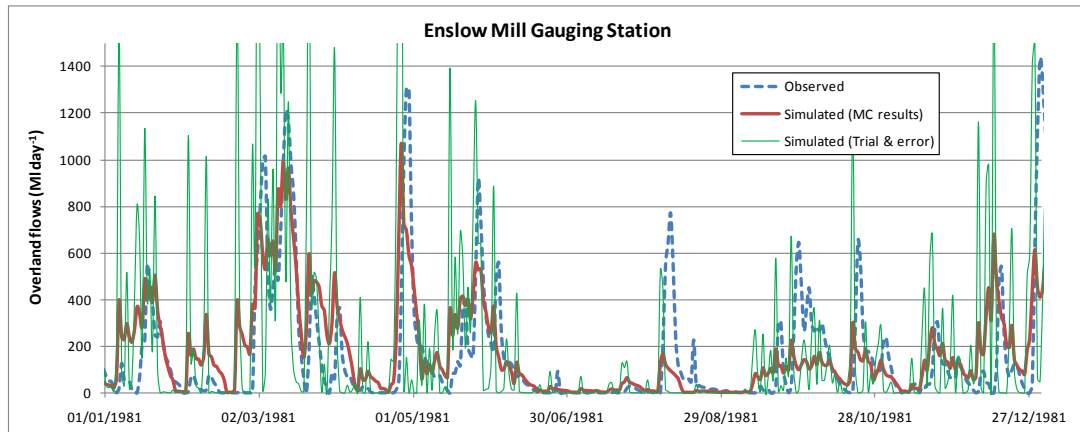


Figure 6 Comparison between the simulated and observed overland flows

Table 1 Baseflow indices and corresponding run-off coefficient values for the geological classes suggested by Bloomfield et al. (2009)

Geological class	Run-off zone	Mean baseflow index	Run-off coefficient
Non aquifer	1	0.243	0.757
Limestone aquifer	2	1.094	0.0
Sandy aquifer	3	0.709	0.291
Non permeable drift	4	0.565	0.435

Table 2 Parameter values of the ARIMA models applied to include the catchment storage

Catchment	P0	P1	P2	P3	P4	P5	P6	P7
Enslow Mill	0.27	0.14	0.14	0.11	0.11	0.09	0.08	0.06
Cassington Mill	0.35	0.15	0.15	0.10	0.10	0.05	0.05	0.05