An assessment of scale issues related to the configuration of the ACRU model for design flood estimation
K. Chetty and J. C. Smithers

ABSTRACT
A continuous simulation modelling approach to design flood estimation has many advantages and overcomes many limitations of commonly used design event approaches. A major concern with continuous simulation using a hydrological model is the scale at which modelling should take place. According to researchers, the level of representation that will preserve the physical chain of hydrological processes, both in terms of scale of representation and level of description of the physical parameters for the modelling process, is a critical question which must be addressed. Objectives of this research were to determine the optimum levels of catchment discretization and soil and land cover information and to assess the optimum use of daily rainfall stations for the configuration of the Agricultural Catchments Research Unit (ACRU) agrohydrological model when used for design flood estimation. Results obtained for selected quaternary catchments in the Thukela catchment and Lions River catchment indicated that modelling at the level of hydrological response units (HRUs), using area-weighted soils information and more than one driver rainfall station where possible, produced the most realistic streamflow volume results when compared with observed streamflows. Design flood estimates from simulated peak flows did not compare well with observed data.

Key words | continuous simulation modelling, design flood estimation

INTRODUCTION
The continuous simulation modelling (CSM) approach to design flood estimation is seen to hold a great deal of potential as inferred by several authors (Hromadka 1987; Rahman et al. 1998; Cameron et al. 1999; Reed 1999). This approach overcomes many of the limitations of the commonly used design event models (Rahman et al. 1998; Cameron et al. 1999). Advantages of CSM for flood estimation are that antecedent moisture conditions are explicitly accounted for in a soil water budget, concepts are more physically based than with other procedures, the effects of complex hydrological systems and river engineered systems can be modelled and a complete hydrograph is obtained, not only peak discharges (Rahman et al. 1998; Cameron et al. 1999; Boughton & Droop 2003).

According to Martina (2004), event-based approaches do not adequately follow the physical chain of hydrological processes in order to assess the role of each of the variables involved in rainfall-runoff processes. However, a CSM approach attempts to preserve the physical relationships between variables. Therefore, in terms of a continuous simulation approach for flood estimation, a critical question which needs to be addressed is the level of representation that will preserve the physical chain of the hydrological processes, both in terms of scale of representation and level of description for the modelling process (Martina 2004).

The transfer of information across scales of both space and time and the associated problems is referred to as the scale issue (Blöschl & Sivapalan 1995). ‘Variability in both space and time is not only an important and significant feature in hydrological science, but one of the many challenges in applied hydrology’ (Woods 2004). The influence
of spatial variability and scale on the hydrologic response of catchments and their importance in hydrological modelling have been widely studied by various researchers (Wood et al. 1988; Blöschl & Sivapalan 1995; Seyfried & Wilcox 1995; Reed et al. 2004; Smith et al. 2004; Booij 2005; Chaplot 2005; Das et al. 2008). In the context of catchment scale hydrological modelling, heterogeneity and variability within catchments pose challenges for dealing with scale issues in hydrology (Blöschl & Sivapalan 1995). Added to this, the use of distributed hydrological models which aim to link conceptualized parameters and processes across scales makes scale issues a very important component of any hydrological modelling applications (Blöschl & Sivapalan 1995).

Various studies have shown improved simulations from more detailed spatial scale and spatial resolution of physical characteristics of a catchment (Seyfried & Wilcox 1995; Wood et al. 1988). However, investigations by Reed et al. (2004) indicate that this type of distributed modelling may not always provide improved simulations. Studies by Loague & Freeze (1985) indicate that ‘there are often problems with quasi-physically based models when used for design flood estimation, especially in terms of scale and the spatial variability of rainfall and soils’. When configuring hydrological models for flood estimation purposes, according to Sangati et al. (2009), the contrasting findings relating to appropriate spatial scales provide little guidance as to the optimum scale of representation for both catchment area and spatially variable catchment properties such as soils and land cover. These contradictory findings indicate the need to further investigate the appropriate scales at which CSM should be applied for improved design flood estimates.

The objective of this research study was to assess the Agricultural Catchments Research Unit (ACRU) agrohydrological model (Schulze 1995) as a tool for design flood estimation based on CSM. This study was the first part of a larger project and focused on assessing scale issues using the ACRU model for design flood estimation. The larger project was aimed at developing a CSM system for design flood estimation. This study entailed investigating in selected catchments the appropriate scale at which continuous simulation should be implemented, the aggregation level of soils and land cover information required to produce optimum simulated streamflow volumes and peaks and the optimum use of daily rainfall stations for the configuration of the ACRU model. The last objective was to compare design flood estimates from peak flows simulated by the ACRU model with observed flood estimates that have been subjected to a frequency analysis. Results from this study would be used to highlight areas that may require investigation and/or modification in the ACRU model when used in a CSM system for design flood estimation.

**METHODS**

The ACRU model Schulze (1995) Version 3.31 was selected for use in this study as the continuous hydrological simulation system. It is a physical–conceptual model developed and tested for southern Africa conditions. The ACRU model operates its soil water budget at a daily time-step in lumped or distributed modes. Soils and land cover information that is used as input to the model can be aggregated at different levels to represent spatial variability within a catchment. In addition, the ACRU model has been used successfully for design flood estimation studies in South Africa (Smithers et al. 1995, 1997; Smithers & Schulze 2000, 2001) and is thus considered suitable to meet the objectives of the study.

The generation of daily streamflow in the ACRU model is based on the sum of the stormflow generated and the baseflow from the catchment in question (Schulze 1995). Stormflow generation is based on a refined version of the US Department of Agriculture (USDA) Soil Conservation Service (SCS) approach. The model also uses SCS techniques for the estimation of peak discharge based on a triangular-shaped unit hydrograph. The daily rainfall input into ACRU is disaggregated within the model into sub-daily time-steps using one of four regionalized synthetic rainfall distributions developed by Weddepohl (1988) and also used for design flood estimation based on adaptations for South Africa to the SCS model (Schmidt & Schulze 1987). The temporal distribution of rainfall events may be influenced by many factors including location, storm duration,
storm depth and season of storm occurrence (Hoang et al. 1999).

In the ACRU model the index of catchment response time or lag $L$ represents the weighted average of the time for stormflow from each point of the catchment to reach the catchment outlet. It is an important factor in determining peak discharge. Lag can be estimated from historical hydrographs or from specific catchment characteristics such as catchment slope, hydraulic length and flow retardance using hydraulic principles or empirical equations (Schulze 1995). The Schmidt–Schulze lag equation is defined:

$$L = \frac{A^{0.35} \text{MAP}^{-1.1}}{41.67y^{0.3} I_{30}^{0.87}}$$

(1)

where $A$ is catchment area (km$^2$), $y$ is average catchment slope (%), MAP is mean annual precipitation (mm) and $I_{30}$ is regional mean of the most intense 30 min period of rainfall. $I_{30}$ can be estimated as the two-year return period 30-min rainfall intensity (mm h$^{-1}$).

Equation (1) was developed using data from small catchments in the US; southern Africa was introduced as an alternative to the original SCS lag equation (Schmidt & Schulze 1987). It incorporates catchment area and mean catchment slope which were determined as dominant physiographic parameters affecting peak discharge. The equation also accounts for climate which influences the soil, vegetation and rainfall patterns, all of which affect the extent to which rainfall enters the soil profile and plays a major role in dominant runoff processes. Climate is represented through the MAP parameter. The two-year return period 30-min rainfall intensity $I_{30}$ was found to affect lag most significantly and was thus incorporated to the equation (Schmidt & Schulze 1987).

**STUDY AREA**

The 29,035.9 km$^2$ Thukela catchment and the neighbouring 353 km$^2$ Lions River catchment located in the KwaZulu-Natal province of South Africa (Figure 1) were selected for use in this research study. The Thukela catchment is divided into 86 so-called quaternary catchments (QCs) while Schulze et al. (2005b) further divided the 86 QCs into 235 subcatchments (Figure 2) to explicitly represent various heterogenic factors which included altitude, soils, topography and vegetation.

In order to meet the objective of this research project, upstream or headwater catchments were selected because the model approach for this study does not include flood routing. Three QCs (QC 6, QC 59 and QC 72) with catchment areas of 129, 152 and 544 km$^2$ respectively were selected as test catchments for this study (Figure 2). This range of catchment area was found to be representative as most of the QCs within the Thukela are less than 500 km$^2$ (Figure 3). However, it was decided that the inclusion of a catchment of approximately 300 km$^2$ would represent a...
better range of test catchment areas. Since no suitable quaternary catchment of this magnitude was found in the Thukela, the 353 km² quaternary U20B (Lions River) of the neighbouring Mgeni catchment was selected.

Rainfall data was obtained from a raster database of daily rainfall developed by Lynch (2004). Point rainfall data recorded at a daily time-step were converted onto a rectangular grid, or raster, using various regression and interpolation techniques. The database consists of quality controlled data from almost 14,000 daily rainfall stations in South Africa. Rainfall stations which fell within and around the Thukela and Lions River catchments were analyzed for their suitability to represent catchment rainfall in the selected catchments. CalcPPTCor, a utility in the ACRU suite of programs that provides an automated method to select the most representative rainfall station for a catchment (termed the ‘driver rainfall station’), was used in this study. Suitable stations are ranked according to an index based on distance of the raingauge from the centroid of the catchment, rainfall record lengths and start and end years of record. The program also automatically calculates the month-by-month precipitation adjustment factors required for each subcatchment. The gauged rainfall data from the selected raingauge are multiplied by the monthly correction factors to ensure that topographical and/or climatological influences within the catchment are accounted for and the input rainfall is representative of the catchment. The program also uses the gridded median monthly rainfall surfaces developed by Dent et al. (1987). The percentage of missing data and infilled data was also taken into consideration when selecting the driver rainfall stations for each subcatchment.

Soils information was obtained from the Institute for Soils Climate and Water land type maps (Land Type Survey Staff 1972–2001) at the 1:250,000 scale which has been translated into ACRU variables by an automated program AUTOSOILS (Pike & Schulze 1995). Land cover information used in this study was obtained from the national land cover database (CSIR 1999). Four major types of land cover were identified in the selected catchments i.e., thicket and bushveld, cultivated crops, forestry and grassland.

The concept of HRUs as introduced by Li et al. (1977) was used in this study. ‘HRUs are distributed, heterogeneously structured areas with common land use and soil or topographic associations controlling their unique hydrological dynamics’ (Flügel 1995). In this study, the HRUs were defined by the four major land cover classes identified and the soils associated with that area. Basically, each HRU is a unique combination of soils and land cover information representing an area homogenous in hydrological response.

In order to investigate the impact of scale issues on model performance it was decided to set up scenarios at three different levels of spatial scale, namely lumped, subcatchment and HRUs as shown in Table 1. In lumped mode, the whole QC was modelled with a single land cover and soil type. At subcatchment scale, each QC was divided into smaller physical subcatchments, as outlined by Schulze et al. (2005a), and each subcatchment had a single land cover and soil type. At the HRU scale, each subcatchment of a QC was divided into four hydrological response units based on the dominant land cover occurring in the QC. For the purpose of this study, the four categories selected were thicket and bushveld, cultivated crops,

![Figure 3](https://iwaponline.com/hr/article-pdf/42/5/401/372569/401.pdf)
forestry and grassland (TCFG). To address the objective of investigating the optimum level of soils and land cover information, each scenario incorporated a different spatial resolution of soils and land cover information. Modal soils or land cover information refers to the soil type or land cover which occurs most frequently in the selected area. Area-weighted soils information was computed by area weighting the soil parameters found in the catchment.

Scenarios were named according to the different levels of land cover and soils information used (Table 2). For example, MA implies Modal land cover and Area-weighted soil and HA implies HRU and area-weighted soil. Scenarios addressing the rainfall stations were referred to as HA1d, where 1d implies that one driver station was used while for scenario HA2d two driver stations were used in the simulations.

The ACRU model was run for the 50-year period of 1950–1999. The scenario which produced the most realistic streamflow volumes when compared with the observed was termed the best simulation and this scenario was then used to estimate peak discharge. The methodology used to estimate design floods was adopted from Smithers et al. (2007). The annual maximum series (AMS) of observed and simulated peak discharge were used in this investigation for the purpose of comparison. According to Smithers et al. (2007), the design floods were estimated by fitting probability distributions to the AMS of peak discharges. L-moments were used to fit the distributions to the AMS. The Log Pearson Type III (LP3) distribution was selected for the estimation of design floods (Alexander 1990).

Table 2 | Levels of scale, soils and land cover information for scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Soils information</th>
<th>Land cover information</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumped</td>
<td>Modal for whole catchment</td>
<td>Modal for whole catchment</td>
<td>Lumped</td>
</tr>
<tr>
<td>MA</td>
<td>Area weighted per subcatchment</td>
<td>Modal per subcatchment</td>
<td>Subcatchments</td>
</tr>
<tr>
<td>HA</td>
<td>Area weighted per subcatchment</td>
<td>Catchment specific HRU</td>
<td>Hydrological response units</td>
</tr>
<tr>
<td>HM</td>
<td>Modal per subcatchment</td>
<td>Catchment specific HRU</td>
<td>Hydrological response units</td>
</tr>
</tbody>
</table>

RESULTS

Accumulated simulated values for all scenarios and accumulated observed runoff are compared for each selected QC (Figure 4) for those periods when non-missing observed data are available and considered to be reliable, as determined from the data analysis undertaken. Judging from a visual inspection of Figure 4, differences between scenarios could be identified. The accumulated totals, the root mean-square error (RMSE) of daily values and the Nash–Sutcliffe efficiency values (NSE) (Nash & Sutcliffe 1970) were however calculated for each scenario for each of the selected QCs.

Considering the total accumulated flow depths, the RMSE values and the Nash–Sutcliffe efficiency values, the ‘best’ scenario was selected. For example, the best scenario selected for QC 59 is Ha3d where the accumulated flow depth was 10,841 mm compared with the observed accumulated flow depth of 11,314 mm, the lowest RMSE = 2.32 mm and the NSE was 0.72 (Table 3). The lumped scenario largely oversimulated the flow depths.

Similarly Tables 4–6 present the results for QCC 6, 72 and U20B respectively, with the best simulations of streamflow depths appearing in bold format.

In addition, a frequency analysis of daily runoff depths for all scenarios for all selected QCs is performed on simulated and observed values which are greater than zero (Figure 5). The model showed reasonable results in QC 59 and QC 72. For QC 6 and U20B there are discrepancies, however. These occurred specifically in the smaller events, which could be due to erroneous observed low-flow measurements or because the low flows were not simulated adequately by the ACRU model.

The scenarios which produced the ‘best’ simulated streamflow depths were then used to simulate peak discharge. A frequency analysis of simulated and observed peak discharge for all selected QCs is illustrated in Figure 6. Figure 7 illustrates that the design floods computed from the simulated peak discharges do not adequately represent the observed peaks with differences in the larger recurrence intervals (1:50 year and 1:100 year) of between 50% and 250%. Fairly realistic results were only obtained for QC U20B.
Table 7 contains the observed and simulated total accumulated streamflow depths and design flood peaks at the 1:50 year and 1:100 year recurrence intervals for the four selected QCs. From the tabulated information, it is evident that for QC 6 there was an overestimation of the design streamflow depths by approximately 4 and 20% at the 1:50 and 1:100 year recurrence intervals, respectively. This overestimation of simulated streamflow depths could be attributed to the driver rainfall station used to represent the rainfall over that catchment having a higher
MAP when compared to a nearby station. However, the design peaks for QC 6 are overestimated by 269 and 278% at the 1:50 year and 1:100 year recurrence intervals, respectively. The overprediction of the design peaks could partially be related to the overestimation of streamflow depth. The oversimulated streamflow volumes were used in the estimation of peak discharge, and could therefore have contributed to the over-simulation of peak discharges. However, such large over-simulation also indicates that there is a significant problem in the estimation of peak discharge in this catchment. This could also be related to inaccurate representation of catchment lag used in the study. The Schmidt–Schulze lag equation used to estimate catchment lag in this study uses MAP as a surrogate for climate and it has already been established that the rainfall station used had a higher MAP than the surrounding stations.

In QC 59, design stormflow depths were under-predicted by approximately 31 and 34%. The design flood peaks were also under-predicted by approximately 23 and 25% for the 1:50 year and 1:100 year recurrence intervals, respectively. These results were better than the results obtained for QC 6. In QC 72, design stormflow depths were under-estimated by 8 and 4% at the 1:50 year and 1:100 year recurrence intervals, respectively. Design peaks were overestimated by 44 and 79% at the 1:50 year and 1:100 year recurrence intervals respectively. As QC 72 was the largest catchment used in the study, it is postulated that the catchment lag could have been a major factor resulting in the overestimation of design floods. The estimated lag could have been too short and hence the resultant peaks too high. The Schmidt–Schulze equation is not suited for use in very large catchments as it was

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Total accumulated streamflow depths and the RMSE of the observed and simulated values for QC 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario</td>
<td>Spatial representation</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Observed</td>
<td></td>
</tr>
<tr>
<td>HA1d</td>
<td>4 HRUs</td>
</tr>
<tr>
<td>HM1d</td>
<td>4 HRUs</td>
</tr>
<tr>
<td>MA1d</td>
<td>1 QC</td>
</tr>
<tr>
<td>Lumped</td>
<td>1 QC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Total accumulated streamflow depths and the RMSE of the observed and simulated values for QC 72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario</td>
<td>Spatial representation</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Observed</td>
<td></td>
</tr>
<tr>
<td>HA1d</td>
<td>8 HRUs</td>
</tr>
<tr>
<td>HA2d</td>
<td>8 HRUs</td>
</tr>
<tr>
<td>HM1d</td>
<td>8 HRUs</td>
</tr>
<tr>
<td>HM2d</td>
<td>8 HRUs</td>
</tr>
<tr>
<td>MA1d</td>
<td>2 subcatchments</td>
</tr>
<tr>
<td>MA2d</td>
<td>2 subcatchments</td>
</tr>
<tr>
<td>Lumped</td>
<td>1 QC</td>
</tr>
</tbody>
</table>
developed on research based in small catchments less than 3 km²; QC 72 is a 544 km² catchment. Parameters used in the estimation of lag such as the 30-minute intensity may no longer hold true on large catchments. Lag is also sensitive to the slope of a catchment, and larger catchments generally imply gentler slopes where stormflow response

Table 6 | Total accumulated streamflow depths and the RMSE of the observed and simulated values for QC U20B

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Spatial representation</th>
<th>Level of land cover</th>
<th>Level of soils information</th>
<th>No. of driver rainfall stations</th>
<th>Total accumulated streamflow (mm)</th>
<th>RMSE (mm)</th>
<th>Nash–Sutcliffe efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6,588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA1d</td>
<td>44 HRUs</td>
<td>Catchment specific</td>
<td>Area weighted</td>
<td>1</td>
<td>3,613</td>
<td>0.8</td>
<td>0.70</td>
</tr>
<tr>
<td>HA5d</td>
<td>44 HRUs</td>
<td>Catchment specific</td>
<td>Area weighted</td>
<td>5</td>
<td>7,105</td>
<td>0.5</td>
<td>0.77</td>
</tr>
<tr>
<td>HM1d</td>
<td>44 HRUs</td>
<td>Catchment specific</td>
<td>Modal</td>
<td>1</td>
<td>4,536</td>
<td>0.83</td>
<td>0.61</td>
</tr>
<tr>
<td>HM5d</td>
<td>44 HRUs</td>
<td>Catchment specific</td>
<td>Modal</td>
<td>5</td>
<td>8,796</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td>MA1d</td>
<td>11 subcatchments</td>
<td>Modal</td>
<td>Area weighted</td>
<td>1</td>
<td>3,527</td>
<td>0.82</td>
<td>0.62</td>
</tr>
<tr>
<td>MA5d</td>
<td>11 subcatchments</td>
<td>Modal</td>
<td>Area weighted</td>
<td>5</td>
<td>7,519</td>
<td>0.51</td>
<td>0.74</td>
</tr>
<tr>
<td>Lumped</td>
<td>1 QC</td>
<td>Modal</td>
<td>Modal</td>
<td>1</td>
<td>15,170</td>
<td>1.65</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Figure 5 | Frequency analyses of simulated and observed daily runoff depths for selected QCs.
may no longer be as rapid as expected on smaller catchments.

In QC U20B, design streamflow depths were over-predicted with percentage differences between simulated and observed of 12 and 16% at the 1:50 and 1:100 year recurrence intervals, respectively. There is a corresponding overestimation of design peaks between 8 and 16% difference at the 1:50 and 1:100 year recurrence intervals, respectively. The results presented in Table 7 indicate that there are problems associated with the estimates of design flood peaks using the ACRU model as configured in this study. QC 59 and QC U20B are the only catchments which yield reasonable results.

**CONCLUSIONS**

Conclusions drawn from the study using the ACRU model in the selected areas for the simulation of streamflow depths can be summarized as follows.

The selected QCs as lumped entities consistently resulted in the large over-simulation of streamflows. Discretizing catchments into subcatchments produced more realistic simulations. However, quaternaries in the Thukela and Lions River catchment should be discretized to the HRU levels since the best simulations were obtained using HRUs.

All scenarios using area-weighted soils information yielded better simulated volumes when compared to the use of modal soils information. It is therefore more appropriate to use area-weighted soils information for the continuous simulation model. Simulated volumes were better when land cover was represented by HRUs than when a single modal land cover was used to represent the land cover in a whole quaternary catchment.

The use of more than one driver rainfall station per sub-QC with appropriate precipitation correction factors yielded better results than modelling using a single driver rainfall station for QC 59 and U20B. This is highlighted in QC 72 and QC 6 where results were not as good since only one rainfall station was used. There were not many reliable representative rainfall stations in the area for QC 72 and QC 6 and the single rainfall station used could not reflect the spatial variations in rainfall within the catchment. The results generally indicate that the better the spatial
variations in rainfall (i.e. represented by more raingauges), the better the simulations.

Comparisons of observed and simulated design flood estimates

The ‘best’ simulations of volume for each of the selected QCs were subsequently used to estimate design peak discharges and these were compared with values computed from the observed flood data. The ACRU model configured at the HRU scale with area-weighted soils information and several suitable driver rainfall stations yielded fairly realistic estimates of the design floods for QC U20B and less for QC 59. The model did not perform well, with differences of 50–250% for the larger recurrence intervals. This poor performance could be attributed to several factors which include inadequate accounting of the temporal distribution of rainfall. The disaggregation of rainfall and selection of

![Figure 7](https://iwaponline.com/hr/article-pdf/42/5/401/372569/401.pdf)
storm distribution types is important, since this type of modelling requires sub-daily increments of rainfall to adequately simulate the peaks.

Another aspect of the model which could result in poor performance in simulating peaks is the estimation of catchment lag. The Schmidt–Schulze lag equation was used in this study and may not have been the most appropriate lag equation for the area. The lag equations themselves may need to be refined by further research when considering the ACRU model for design flood estimation.

**ACKNOWLEDGEMENTS**

This research formed part of a Water Research Commission Project (K5/1318) entitled ‘The Development of a Continuous Simulation System for Design Flood Estimation in South Africa’. We therefore wish to thank the WRC for funding this project and the members of the steering committee for their guidance during the project.

**REFERENCES**


Chaplot, V. 2005 Impact of DEM mesh size and soil map scale on SWAT runoff, sediment and NO3-N loads predictions. *Journal of Hydrology* 312, 207–222.


Land Type Survey Staff 1972–2001 *Land Types of South Africa on 1:250,000 Scale*. ARC-ISCW, Pretoria, South Africa.


First received 16 December 2009; accepted in revised form 21 September 2010. Available online June 2011