Introduction

Although there is uncertainty in terms of rate of change, the earth’s climate is undoubtedly being perturbed through global warming. This is confirmed by past records showing the close correlation between the temperature changes (globally, an increase of 0.5 deg. C over the last century) and altering $CO_2$ concentration. To be more specific, ultimate temperature increases on land seem likely to be within the range of 3-6 deg. C in winter and 2-4 deg. C in summer (Houghton et al., 1996). The main cause of climate change is the production of greenhouse gases (released into the atmosphere through the use of fossil fuels), of which carbon dioxide constitutes at least half. This phenomenon has been accelerated by deforestation (particularly tropical) and rangelands overgrazing (mainly temperate), although these effects are not comparable with the increase due to fossil fuels in this century.

Various definitions may be given for climate, climate variability and climate change. However, the distinction between “variability” and “change” is clear in principle, though it is not easy to apply in practice. Generally, the term ‘climate change’ is used when there is significant long-term change in the mean of a climate variable, whereas, ‘climate variability’ refers to natural variation from year to year.

General Circulation Models (GCMs) are in principle the most appropriate tools for predicting climate change as they provide estimates of the hydro-climatological balance. However, because these estimates are integrated over spatial gridsquares of some 300km, they are not appropriate for use on a smaller, catchment scale. This is also true for the temporal scale, since although GCMs operate in short time steps, the rainfall intensities are estimated uniformly over a large spatial scale (see e.g. Osborn, 1996, 1998). GCM results can therefore at most provide only general tendencies, and their usefulness is limited when compared to the...
Various approaches have been developed to study the response of catchments to the climate change in order to extract hydrologic information at the catchment scale through rainfall data. An example is devising a variety of nesting schemes from GCM results such as Limited Area Meteorological (LAM) models, Macroscale Hydrologic Models (MHM), or sub-grid parameterisations (Loaiciga, et al., 1996).

This nesting approach uses a much higher resolution model nested within a GCM. For example, a small region such as Western Europe is selected and a Regional Climate Model (RCM), with grid size of 50 km or smaller is run (e.g. McGregor et al., 1993 in Australia). However, the RCM is still conditioned by the GCM boundary conditions, and computational expense dictates that simulations cannot be performed over long periods.

In another approach to the sub-grid scale method used by Wigley et al. (1990), the downscaling approach is based on statistical linkages between the local and large scale climate (von Storch, et al., 1993). The main assumption here is that orographic and geographic (land/sea contrast) factors are responsible for local scale variations. The effects of these factors can then be estimated by regression relationships between local and large-scale climate parameters. The statistical linkages are assumed to remain valid as far as the future climate is concerned, and so future hydrological scenarios for any spatial scale can be predicted by using GCM scenarios and the regression relationships obtained from observed data. This procedure has an advantage over the nested modelling approach due to the ease with which various methods can be tested and the speed with which long simulations may be generated.

Future climate prediction is subject to many uncertainties as far as policy responses in controlling the emission of CO2 in all parts of the world are concerned. Numerous scenarios have been developed based on various annual rates of CO2 increase up to the year 2100 (Houghton et al., 1996). In recent years, efforts have been made to construct catchment scale climate scenarios using downscaling methodologies employing GCM outputs such as weather circulation indices and temperature rather than using rainfall directly from GCM outputs (Kilsby et al., 1998). Therefore, a statistical downscaling approach constitutes the main element of the scenario construction methodology used for impact assessment. With this approach, time series of sufficient duration to represent long-term variability, including extreme events and droughts, which control the reliability of water resource systems, can be generated. This approach to scenario construction will be demonstrated here at the daily time-scale and catchment space-scale.

The objective in this paper is to assess the impacts of climate change on generated daily rainfall series for a small catchment. That is to generate and validate rainfall sequences using an stochastic model, which can account for climate change using modified parameters; for present/control conditions of climate. Then, to generate control and future climate scenarios using the stochastic model of rainfall fitted above. Sequences of rainfall can then be routed to a rainfall-runoff model for water resources impact assessments (this is not investigated here). The results obtained for future scenarios are compared with corresponding values for the control scenarios.
**Catchment and Available Data**

**General Catchment Description**

Kassilian catchment in the north of Iran was adopted as the focus of this study. The study catchment has minimal human influence (no reservoirs) and relatively accurate data. Figure 1 shows a map of the catchment and the main drainage network. It has an area of approximately 66.75 square kilometres at Valikbon outlet. The average annual rainfall for the period of 1976-88 is around 844 millimetres. The catchment has a humid temperate climate with precipitation evenly distributed throughout the year. The contribution of snowmelt to precipitation for the study catchment is generally insignificant.

**Data Sets**

Historic series of daily rainfall was available for a period of twelve years (1/10/1976-31/9/1988). Daily rainfall data were available as weighted mean of 5 stations distributed over the area of the catchment. Statistical information for the study catchment is summarised in Table 1.

The justification for using data for the this period for the study is as follows: the apparent increase in variability and extreme events in next years has provided some doubts for climate change in more recent years. Therefore, the 1990s and later may be more affected by global warming than the previous years (pre-1988 period). In order
not to take the risk of the existence of the recent probable climate change in the observed data, it was considered preferable to use the earlier record (previous-1988) for analysis.

Methodology Used for Impact Assessment

Compare to various daily rainfall models used in literature for fitting to observed rainfall data and then simulation, the selected model for this study should be able to fulfil two objectives. 1- The model should be able to reflect physical characteristics of atmosphere relatively which affects weather moisture and then rainfall conditions. 2- The model structure should be able to accommodate a downscaling approach in order to use regional GCM data to a catchment scale as small as Kassilian. Regarding the 1st objective, there is an extensive literature on daily and shorter period rainfall modelling (for example Foufoula-Georgiou and Krajewski (1995)). For daily rainfall, two broad categories of model have been employed: discrete time series models and point process models. The former uses a discrete time increment (e.g. a day or an hour) whilst the point process models use a continuous time model for the occurrence of rainfall events, which characterises indirectly the wet and dry periods, and to describe the random amounts of rainfall associated with the wet periods. Regarding the former, a typical category of models is ARMA models. Of the point process models, the Neyman-Scott models (White Noise or Rectangular Pulses) have shown to be at least as good as other models (Wilby, 2001). These models, compare to the well-known rainfall models such as Markov chain models, can therefore, reflect relatively physical characteristics of atmosphere which affects weather moisture and then rainfall conditions. This provides a context for accommodating those physical characteristics that can then be used for reflecting the change of rainfall in future condition. As regards to the 2nd objective, the downscaling approach uses regression relationships between atmospheric circulation indices (ACIs) and rainfall statistics. The relationships are then used to predict the rainfall statistics for future conditions using GCM outputs. Having selected the rainfall model, the methodology is summarised as:

(a) to fit the selected model to observed data (that present climate condition),

(b) to use the fitted model to generate hourly rainfall data using a random generator scheme and then aggregated to daily values,  

(c) to validate the generated data against historic values employing a 2 stage scheme,

(d) Having validated for present condition, to use the model to generate future rainfall data employing the downscaling approach explained in below.

The Neyman-Scott Rectangular Pulses (NSRP) model has been used for this study.

Table 1. Data summary for rainfall data (1976-88) consisting of means, standard deviations at daily and annual aggregation levels

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Mean (mm)</th>
<th>Std (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>2.31</td>
<td>5.20</td>
</tr>
<tr>
<td>Annual</td>
<td>844.0</td>
<td>144.54</td>
</tr>
</tbody>
</table>
(Rodriguez-Iturbe et al., 1987; Cowpertwait, 1991a; Cowpertwait et al., 1996) for the following reasons:

(a) It has a realistic physical structure. This makes the interpretation of the parameters of the model easier,
(b) It preserves historic rainfall statistics at various levels of aggregation (hourly and above),
(c) It requires only 5 parameters to be estimated. This simplifies the parameter estimation procedure, and
(d) Its parameters can easily be re-estimated in a perturbed climate by using the relevant perturbed statistics, which are functions of GCM outputs.

Rainfall Modelling for Present Climates

Model Calibration

The NSRP model uses a clustered process for the arrival of rain-cells (Cowpertwait, 1991a). The parameters are: $1/\lambda = \text{mean time between storm origins (hour)}$, $1/\beta = \text{mean waiting time for cells after the storm origin (hour)}$, $\nu = \text{mean number of rain cells per storm}$, $1/\eta = \text{mean cell duration (hour)}$, and $1/\zeta = \text{mean cell intensity (mm/hour)}$. To deal with seasonality, the NSRP parameters are estimated separately for each month. The fitting procedure, based on daily rainfall, involves estimating the five parameters of the NSRP model through minimising the following sum of squares function:

$$ S = \sum_{i=1}^{m} w_i \left( 1 - \frac{f_i}{\hat{f}_i} \right)^2 $$

where $\lambda, \beta, \eta, \nu, \zeta > 0$, and $\nu > 1$; $f_i = f_i(\lambda, \beta, \eta, \nu, \zeta)$ is a model function that defines a particular model statistic (see Rodriguez-Iturbe et al. (1987) and Cowpertwait (1991a)) and $\hat{f}_i$ is its estimated statistic from the historical rainfall data; $m$ refers to the number of moments which is taken to be equal to or greater than the number of parameters. The weights ($w_i$) allow greater weight to be given in fitting some statistics. Here an arbitrary value of $w_i=10$ is chosen for the term relating to the mean daily rainfall and value of $w_i=1$ is applied to the remaining statistics. In practice, finding a value close to zero for $S$ is the aim.

The average values of the following statistics of the observed 12 years of continuous daily rainfall time series, provided as the average of a number of point values for the catchment, were derived for each calendar month:

(i) daily mean, (M24),
(ii) daily variance, (V24),
(iii) proportion of dry days, (PD),
(iv) proportion of dry days preceded by a dry day, (PDD), and
(v) proportion of wet days preceded by a wet day, (PWW).

The model parameters were then fitted using a quasi-Newton algorithm for finding the minimum of the objective function, subject to fixed upper and lower bounds on the independent parameters.

Data Generation and Validation

The model parameter estimates (60 parameter estimates in total, 5 for each month) were considered to be physically realistic. The fitted model was then used to generate rainfall data which reproduced key daily rainfall statistics (mean, variance, probability of dry days, wet and dry
transition probabilities). The generated data were validated using a two stage approach explained below.

1st validation stage

Statistical resemblance to the daily historic rainfall record is checked. That is, preserving statistics which were implicitly designed for (i.e. mean, variance, autocorrelation and skewness coefficient of daily rainfall). In this respect, the basic statistics were in reasonable agreement. In addition to that, the approach should produce sequences which reflected the observed structure at various aggregation levels e.g. monthly and annual. This also more or less has been fulfilled. Therefore, it was considered that the generated rainfall data were in satisfactory agreement with the observed data (see Table 3 as historic and present climate lines to be compared).

2nd validation stage

As indicated, the model reproduced the five chosen statistics fairly well, but the reproduction of other statistics is not usually guaranteed. Testing the model’s ability to reproduce rainfall properties not used in the fitting procedure, but of practical importance, is a necessity for the final stage of validation. For example, in drought situations, a reproduction of the lengths of dry spells and of their frequencies may be regarded as important. So dry spell lengths can be employed as a guide in the related decision-making. In view of the intended use of the rainfall model, which can be for water resources impact studies associated with climate change, a validation test was carried out for this stage. That is, the frequencies of occurrences of dry days or the distribution of the number of consecutive dry spell sequences were calculated and compared. In this respect, a year was divided into four seasons: January, February and March; April, May and June; July, August, and September; and October, November and December. The number of occurrences of a sequence of n dry days is found for both 25 generated series and the historic series over the 12 year period. That is, the number of times that only one day separated rainy days is counted and this process is repeated for two days, three days and so forth until the longest dry run is counted. The results of this comparison, between the historic values and the corresponding statistics for 25 sequences (in terms of, means, mean ± standard deviation, and mean - standard deviation) are presented in Figure 2. It was found for season 3 that the model reproduced the historical dry spell sequences well, except for very short dry sequences, which were over-estimated. Moreover, the mean number for the very long dry sequences in the generated series is somewhat underestimated. The longest historic dry spell sequence is also 23 days; while longer spells were observed in the generated series.

The overall results showed the generated rainfall data reflected the corresponding historic statistics satisfactorily and the rainfall model could be adopted with sufficient confidence to assess the impacts of climate change for the study catchment.

Scenario construction

Downscaling approach

Regression relationships between atmospheric circulation indices (ACIs) and two rainfall statistics, monthly-mean daily rainfall (MDR) and proportion of dry days
(PD) were taken as the basis for downscaling GCM outputs to the catchment scale and then for re-estimating the NSRP model parameters. Additionally the 24 hour variance (VAR24) was used in re-estimation, obtained as described below. Although it would be desirable to use downscaled estimates of other statistics, this is not currently possible (e.g. Murphy, 2000). However, MDR and PD are considered able to capture the most important changes.

The regression approach of Kilsby et al. (1998) was used to estimate the statistics for both control (CON) and perturbed (SUL) rainfall transient scenarios using Hadley Centre GCM outputs of atmospheric variables with the cooling effects of aerosols included (Mitchell et al., 1995). The predictor variables used are three independent ACIs such as mean pressure (P), zonal flows (U) and meridional flow (V) all derived from observed or GCM grids of mean sea level pressure (MSLP). The following relations for the expected values of MDR and PD at site i, year j, and month k, denoted MDRijk, and PDijk, respectively, were established:

\[
E(MDR_{ijk}) = r_m \exp\{\alpha_p + \alpha_u U_{jk} + \alpha_v V_{jk} + \alpha_p P_{jk}\}
\]  
\[
E(PD_{ijk}) = r_p \left[1 + \exp\{-\beta_u U_{jk} - \beta_v V_{jk} - \beta_p P_{jk}\}\right]
\]

where \(r_m\), in Equation 2, is a correction ratio to allow for the bias resulting from retransformation from \(\ln(MDR)\) to MDR in model fitting, and \(r_p\), in Equation 3, is again a correction ratio for the retransformation bias.

Values of the regression coefficients (U, V, and P) and correction ratios (r), together with the coefficients of determination (\(R^2\)), derived using ACI, site and point rainfall data
for the region, are given in Table 2.

The regression relations were then used to estimate MDR and PD values using the HADCM2 GCM outputs for control (CON) and perturbed future (SUL) scenarios. The results for CON and SUL conditions in addition to the observed statistics usually show that some discrepancies between the GCM control rainfall statistics and the observed rainfall statistics are evident due to the inability of the GCM to reproduce regional climate patterns accurately (for details see Kilsby et al., 1998). To avoid these discrepancies in deriving future (FUT) statistics, a standardisation scheme was adopted to incorporate the changes to observed values using the statistics of the CON and SUL GCM experiments. That is, on a monthly basis, relative changes of the two rainfall statistics, denoted FUTMDR and FUTPD, were estimated by multiplying the ratio of the statistic of the perturbed (SUL) GCM scenario and the statistic of the control (CON) GCM scenario by the corresponding observed statistic:

\[
FUT = \frac{SUL}{CON} \times Observed
\]

The assumptions in the above approach are that the present day frequency of weather types and circulation patterns will be changed as a result of climate change, but that each weather type will retain its associated precipitation characteristics.

**Rainfall Generation for Future Condition**

Future rainfall statistics, denoted FUT(GCM), were obtained as described above, but using the corresponding catchment historic rainfall statistics. This approach is particularly useful when the discrepancies between the historic data and control scenarios are significant. In this respect, these discrepancies are considered to have come from two sources:

(i) the inability of the GCM to reproduce satisfactorily the behaviour of the present climate and (ii) the relationships derived for predicting MDR and PD from atmospheric and physiographic variables relate to rainfall at a point, whereas the historic statistics for the study catchment relates to average catchment rainfall.

The parameters of the NSRP model were re-estimated using the predicted (FUT) values of MDR and PD calculated from above equation as well as the variance (VAR). Regarding the VAR, it was assumed that VAR would increase under future conditions as the mean increases, i.e. that the coefficient of variation (CV) will remain constant. The other possible assumption is that VAR would remain constant, i.e. CV would be dependent on mean. For a strict choice between these

### Table 2. Values of regression coefficients and goodness of fit for MDR and PD

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Constants</th>
<th>( U )</th>
<th>( V )</th>
<th>( P )</th>
<th>( r )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDR</td>
<td>73.5</td>
<td>0.0051</td>
<td>-0.00418</td>
<td>-0.80230</td>
<td>1.130</td>
<td>0.33</td>
</tr>
<tr>
<td>PD</td>
<td>-96.1</td>
<td>-0.0078</td>
<td>-0.00214</td>
<td>0.11265</td>
<td>0.787</td>
<td>0.46</td>
</tr>
</tbody>
</table>
assumptions, there is insufficient information currently available and, therefore, this is an area subject to further research. Here, the statistics of the variance (VAR) for future GCM conditions, i.e. FUT(GCM), were obtained for each month using the statistics of MDR for future GCM conditions and the CVs of present conditions (historic values) as follows.

\[
\text{VAR}_i^{\text{FUT(GCM)}} = (\text{CV}_i^{\text{HIS}} \times \text{MDR}_i^{\text{FUT(GCM)}})^2
\]

where, \(i = 1, 2, ..., 12\).  

Two NSRP parameters were then re-estimated; \(\mu\) (the mean cell intensity) and \(l\), (the rate of storm arrival) using MDR, PD and VAR; the other three parameters (b, n, and h) were assumed to remain constant in the future climate and remained fixed in the re-estimation procedure. The two re-estimated parameters are the most appropriate, since there is a direct relationship between the parameters and these statistics defined in the rainfall model structure (for the analytical relationships between the statistics and parameters see Rodriguez-Iturbe et al. (1987) and Cowpertwait (1991a)). Furthermore, since only small changes are predicted in the NSRP model parameters, and no major changes are evident in the dominant precipitation mechanism (Kilsby et al., 1998), the assumption that some of the NSRP model parameter estimates remain constant is reasonable.

Results and Discussions

The NSRP model with the re-estimated (FUT) parameters was used to generate an ensemble of 25 synthetic series of daily rainfall data, each of the same length as the historic series (12 years). The ensemble average MDR and PD values, i.e. FUT(SIM), are shown in Figure 3 together with MDR and PD of historical values.

The statistics of MDR when compared with the historic values indicate an significant increase in January and March (months 4, and 6, respectively). The results for MDR in this figure follow more or less the pattern (but not the magnitude) of the historic statistics. The statistic of PD again follows the pattern of the historic one and not the magnitude, as significant decreases are evident here for 2 or three months. Moreover, results for a number of statistics such as mean, variance, lag-one autocorrelation coefficient (L1Acc), and skewness coefficient (SC) of ensembles of 25 generated series for present and future climates are shown in Table 3 together with the historic counterparts at daily, aggregated monthly, and aggregated annual levels. In this table, generally, the mean values show an overall increase in rainfall amount of around 6% mainly in January and March (see Fig. 3). This is moreover true for the monthly and annual aggregation level. However, significant increases in daily as well as monthly variances are noted. The increase in annual variance is around 28%.

The results suggest wetter conditions in the study catchment, particularly during the winter season, consistent with the results of other studies performed elsewhere (Hulme and Jenkins, 1998). The rainfall regime is also expected to be more variable.

Conclusion

A Neyman-Scott Rectangular Pulses model has been applied to areal average of daily rainfall data for Kassilian catchment. A model fitting procedure has been employed:
one based on the use of statistics such as daily mean, daily variance, proportion of dry days, proportion of dry days given the previous day dry, and proportion of wet days given the previous day wet, denoted as M24, V24, PD, PDD, and PWW, respectively. PD and PDD statistics were selected in this study due to their potential capability for low flow simulations as they use daily transition probabilities within the fitting procedure. That is because, they give an appropriate model fit to the historic dry spell sequences as far as the validation procedure is concerned. The conclusions are summarised as: 1- When dealing with NSRP modelling for rainfall generation, various fitting schemes may be chosen with various fitting results obtained. However, the appropriate rainfall modelling and fitting scheme should be selected on the basis of the intended application of generated series; the selected scheme may not be as good as other schemes as far as discrepancies between standard historic and fitted statistics are concerned, but should perform satisfactorily in reproducing the rainfall characteristics to

![Figure 3. Summary statistics (MDR and PD) for 25 simulated sequences (FUT(SIM)) compared to historic (HIS) for Kassilian catchment rainfall (Oct.=Month 1)](image)
which a water resource system is sensitive. 2- A rainfall model fitting procedure which uses wet and dry period transition probabilities is an appropriate candidate on the grounds of its potential capability for low flow simulation in climate impact assessment studies. 3- The overall results obtained showed rainfall data are projected to be increased with higher magnitude of the winter season. 4- A cascade of uncertainties from emissions, GCM, downscaling to small spatial scales exist that should be addressed in further studies before using generated rainfall data as input to a system catchment rainfall-runoff model for generating streamflow values in water resources impact assessment.

Acknowledgement
Data from the Hadley Centre were provided by the Climatic Research Unit at the University of East Anglia.

References


[3]. Foufoula-Georgiou, E., and Krajewski, W., 1995 Recent advances in rainfall

<table>
<thead>
<tr>
<th>Level</th>
<th>Type of data, statistics, and climate scenario</th>
<th>Mean (mm)</th>
<th>Variance (mm)^2</th>
<th>L1Acc</th>
<th>SC.</th>
</tr>
</thead>
<tbody>
<tr>
<td>daily</td>
<td>Historic</td>
<td>2.31</td>
<td>27.04</td>
<td>0.241</td>
<td>4.441</td>
</tr>
<tr>
<td></td>
<td>generated sequences: present</td>
<td>2.29 (0.004)</td>
<td>25.85 (0.781)</td>
<td>0.237</td>
<td>3.885</td>
</tr>
<tr>
<td></td>
<td>mean (variance)</td>
<td>2.42 (0.002)</td>
<td>33.32 (1.122)</td>
<td>0.120</td>
<td>3.122</td>
</tr>
<tr>
<td></td>
<td>future</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>monthly</td>
<td>Historic</td>
<td>2.30</td>
<td>1.96</td>
<td>0.064</td>
<td>1.493</td>
</tr>
<tr>
<td></td>
<td>generated sequences: present</td>
<td>2.28 (0.003)</td>
<td>1.92 (0.024)</td>
<td>0.076</td>
<td>1.095</td>
</tr>
<tr>
<td></td>
<td>mean (variance)</td>
<td>2.40 (0.002)</td>
<td>2.77 (0.322)</td>
<td>-0.002</td>
<td>1.112</td>
</tr>
<tr>
<td></td>
<td>future</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>annual</td>
<td>Historic</td>
<td>2.31</td>
<td>0.157</td>
<td>0.009</td>
<td>-0.103</td>
</tr>
<tr>
<td></td>
<td>generated sequences: present</td>
<td>2.29 (0.003)</td>
<td>0.165 (0.001)</td>
<td>0.320</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>mean (variance)</td>
<td>2.42 (0.001)</td>
<td>0.212 (0.111)</td>
<td>0.001</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>future</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Comparison of basic statistical characteristics (mean, variance, lag-one autocorrelation coefficient (L1Acc), and skewness coefficient (SC)) of historic and the corresponding 25 generated series of daily rainfall data for both present and future climates, at daily, monthly and annual levels.


