

using science to create a better place

Continuous Estimation of River Flows (CERF)

Science Report SC030240

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned collaboratively and funded on a shared-cost basis by the Centre for Ecology and Hydrology and the Environment Agency's Science Programme.

Published by:

Environment Agency, Rio House, Waterside Drive,
Aztec West, Almondsbury, Bristol, BS32 4UD
Tel: 01454 624400 Fax: 01454 624409
www.environment-agency.gov.uk

ISBN: 978-1-84432-874-1

© Environment Agency April 2008

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.

The views and statements expressed in this report are those of the author alone. The views or statements expressed in this publication do not necessarily represent the views of the Environment Agency and the Environment Agency cannot accept any responsibility for such views or statements.

This report is printed on Cyclus Print, a 100% recycled stock, which is 100% post consumer waste and is totally chlorine free. Water used is treated and in most cases returned to source in better condition than removed.

Further copies of this report are available from:
The Environment Agency's National Customer Contact Centre by emailing:
enquiries@environment-agency.gov.uk
or by telephoning 08708 506506.

Authors:

Griffiths, J., Keller, V., Morris, D., Young, A.R.

Dissemination Status:

Release to all regions
Publicly available

Keywords:

Rainfall-runoff modelling, ungauged basins, time series, semi-distributed model, calibration, parameterisation

Research Contractor:

Centre for Ecology and Hydrology (Wallingford)

Environment Agency's Project Manager:

Stuart Allen, Science Department

Science Project Number:

SC030240

Product Code:

SCH0308BNVZ-E-P

Science at the Environment Agency

Science underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us and helps us to develop monitoring tools and techniques to manage our environment as efficiently and effectively as possible.

The work of the Environment Agency's Science Department is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The science programme focuses on five main areas of activity:

- **Setting the agenda**, by identifying where strategic science can inform our evidence-based policies, advisory and regulatory roles;
- **Funding science**, by supporting programmes, projects and people in response to long-term strategic needs, medium-term policy priorities and shorter-term operational requirements;
- **Managing science**, by ensuring that our programmes and projects are fit for purpose and executed according to international scientific standards;
- **Carrying out science**, by undertaking research – either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.



Steve Killeen

Head of Science

Executive summary

The Environment Agency has a statutory duty to contribute to the sustainable development of water resources. The Environment Agency's Catchment Abstraction Management Strategies (CAMS), implementation of the Habitats Directive and the Water Framework Directive (WFD), and routine abstraction licensing and discharge consenting, all require the Environment Agency to estimate water resources potential for every catchment in England and Wales. This necessitates the estimation of river flows at both gauged and ungauged locations. For some of these activities, such as the setting of discharge consents and licensing of small abstractions, it is sufficient to encapsulate this information using a statistical description of the flow regime such as the flow duration curve. However, for the remaining activities, the use of a time series of river flows is either required, or would be advantageous. The CERF (Continuous Estimation of River Flows) project has been focused on addressing that need.

The project was funded on a shared-cost basis by the Environment Agency and the Centre for Ecology and Hydrology (CEH). From the Environment Agency's perspective the overarching objective of this project was to develop a methodological framework and suite of tools for continuous estimation of daily time series of flows for ungauged catchments. This framework and tool set would underpin any subsequent software development for the Environment Agency desktop server system for use by operational staff. The overarching objective of the CEH was to improve the knowledge base underpinning the estimation of river flows within ungauged catchments. The CERF project contributes to answering specific science questions identified within the five-year CEH science strategy, 'Health and Wealth of the Environment'.

Within the CERF model the generic concepts of Hydrological Response Units (HRUs) are used to define a flexible model structure in which catchment descriptors of vegetation, soil type, topography and geology are used to define relatively complex, unique model structures in each catchment. The number of HRUs depends on the complexity of the catchment. So, for example, small catchments with relatively similar soils, geology and vegetation will have relatively few, while large diverse catchments may have many. The model consists of three basic HRU components;

- a soil moisture accounting HRU based upon the FAO56 soil moisture model coupled to a generalised interception model;
- a semi-distributed quick flow routing scheme that explicitly accounts for flow paths through a catchment;
- a lumped linear reservoir routing scheme for the routing of slow flow to the catchment outlet.

These components come together to form a regionalised rainfall–runoff model that can predict daily time series of natural river flows for ungauged catchments. The model is complementary, rather than a replacement methodology for the existing Low Flows 2000 model for predicting flow duration statistics. The CERF model has been extensively validated during the project and has been shown to give results that are stable when applied to independent catchments and that are generally of significantly greater accuracy than current approaches based on the transposition of flow time series from a suitable non-nested analogue catchment. Recommendations are made on when to use the CERF model and when to use transposed data from analogue catchments.

The application of a rainfall–runoff model within an ungauged catchment provides the user with the challenge that the quality of the model output cannot be evaluated through comparison with measured flow data.

To address this the performance of the CERF model has been extensively evaluated to identify the causes of poor model fit (e.g. poor input precipitation data) to provide guidance on when CERF is likely to produce good simulations and when the simulations might be compromised by poor quality data or limitations of the modelling approach.

It is the longer term aspiration of the Environment Agency to provide CERF to Environment Agency users through an appropriate IT solution. In the interim a bureau service is available by which streamflow simulations, and secondary products such as rainfall grids, can be requested for any river reach within England and Wales.

Contents

1	Introduction	1
2	Background to the project and requirements	3
3	Summary of the CERF model	6
4	Task 1.1: Estimation of precipitation inputs	11
5	Task 1.2: Model scheme for representing evapotranspiration processes	13
6	Task 1.3: HRU model scheme for representing rainfall interception and soil moisture	15
7	Task 1.4: Model scheme for routing effective precipitation to streamflows	17
8	Task 2.1: An evaluation of the sources of predictive uncertainty within the CERF model	22
9	Task 2.2: Guidance for prediction method selection based on minimisation of uncertainty	25
10	Task 2.3: Time-series indicators of operational utility	27
	References	30
	List of abbreviations	31

Figures

Figure 1.1	Schematic diagram of inputs and outputs of CERF modelling scheme	2
Figure 3.1	Flow diagram illustrating the three main stages of the CERF modelling process: estimation of evapotranspiration, soil moisture modelling scheme and semi-distributed routing model.	7
Figure 3.2	Soil moisture accounting structure.	9
Figure 5.1	Calculation of actual evapotranspiration for different crop types from PE estimates and land-cover parameters.	14
Figure 6.1	Use of LCM2000 and HOST data, and subsequent calibration and validation of the soil moisture model using data from 357 identified test catchments.	16
Figure 7.1	Flowchart of semi-distributed structure approach.	18
Figure 7.2	Simulated and observed hydrographs for the example catchments.	20
Figure 7.3	Location of example catchments.	21

Tables

Table 1.1	Technical reports.	1
Table 7.1	Performance comparison between the semi-distributed and the lumped routing scheme over the calibration period within the common catchments used for calibration (138 catchments).	19
Table 7.2	Performance comparison between the semi-distributed and the lumped routing scheme over the calibration period within a set of independent evaluation catchments (70 catchments).	19
Table 8.1	A classification of catchment depending on their performance level.	23

1 Introduction

The Continuous Estimation of River Flows (CERF) project has delivered a regionalised rainfall–runoff model for the prediction of time series of river flows (CERF model). The CERF model compliments the existing Low Flows 2000 model which predicts flow duration statistics. The CERF model has been extensively validated during the project and has been shown to give results that are stable and that are generally of significantly greater accuracy than current approaches based on the transposition of flow time series from adjacent gauged catchments. The project makes recommendations on when to use the CERF model and when to transpose data from analogue catchments.

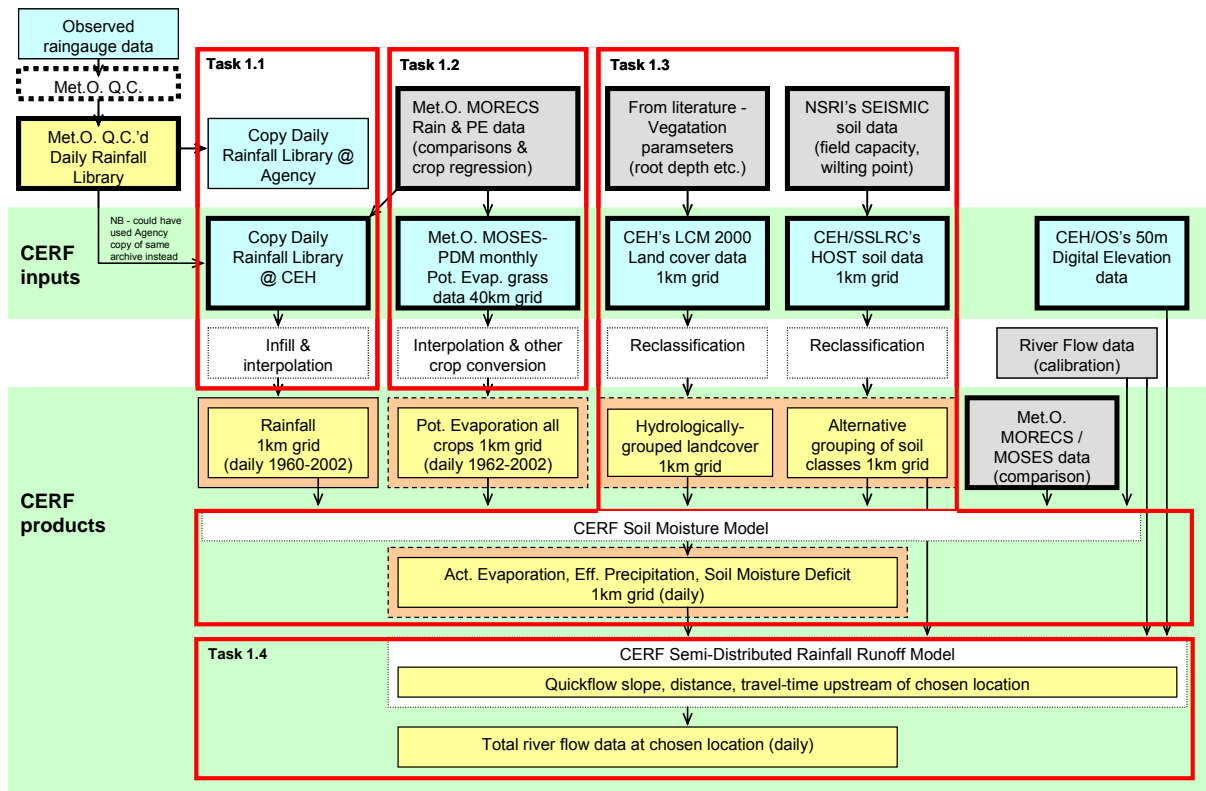
The application of a rainfall–runoff model within an ungauged catchment provides the user with the challenge that the quality of the model output cannot be evaluated through comparison with measured flow data. To address this the performance of the CERF model has been extensively evaluated. Causes of poor model fit (e.g. poor input precipitation data) are identified and guidance provided on when CERF is likely to produce good simulations and when the simulations might be compromised by poor quality data or limitations of the modelling approach.

The background to the project and requirements from the project are discussed within Section 2.

The project was structured into two phases; each phase was sub-divided into tasks, activities and deliverables. Each task was written up as a technical report (see Table 1.1). An overall summary of the CERF model is given in Section 3. Greater detail of the activities and deliverables of each task within the model design and calibration phase is then given in Sections 4 to 7. Assessment of model performance and recommendations for implementation are presented in Sections 8 to 10. Figure 1.1 illustrates the division of the modelling phase into tasks (1.1 to 1.4). The flow of information and data required for the modelling process is shown in relation to model outputs and products from the project.

Table 1.1 Technical reports

Task	Technical report title
1.1	Estimation of precipitation inputs
1.2	Model scheme for representing evapotranspiration processes
1.3	HRU model scheme for representing rainfall interception and soil moisture
1.4	Model scheme for routing effective precipitation to streamflows
2.1	An evaluation of the sources of predictive uncertainty within the CERF model
2.2	Guidance for prediction method selection based on minimisation of uncertainty
2.3	Time-series indicators of operational utility



KEY to acronyms

CEH = Centre for Ecology & Hydrology
 Met.O. = Meteorological Office
 UK LCM2000 = UK Land Cover Map 2000
 NERC = Natural Environment Research Council (includes CEH - Centre for Ecology & Hydrology)
 HOST = Hydrology Of Soil Types
 NSRI = National Soil Research Institute
 SEISMIC = Spatial Environmental Information System for Modelling the Impact of Chemicals
 SSLRC = Soil Survey & Land Research Centre
 OS = Ordnance Survey

KEY to boxes

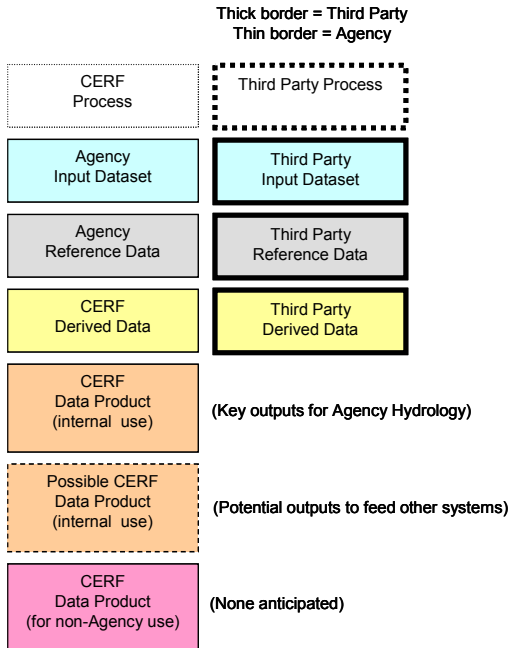


Figure 1.1 Schematic diagram of inputs and outputs of CERF modelling scheme

2 Background to the project and requirements

This project has been funded on a shared-cost basis by the Environment Agency and the Centre for Ecology and Hydrology (CEH). From an Environment Agency perspective the overarching objective of this project has been to develop a methodological framework and suite of tools for continuous estimation of daily time series of flows for ungauged catchments. This framework and tool set would underpin any subsequent software development for the Environment Agency desktop server system for use by its staff.

The overarching objective of the project from the perspective of the CEH was to improve the knowledge base underpinning the estimation of river flows within ungauged catchments. The CERF (Continuous Estimation of River Flows) project contributes to answering specific science questions identified within the five-year CEH science strategy, 'Health and Wealth of the Environment'.

The Environment Agency has a statutory duty to contribute to the sustainable development of water resources. The Environment Agency's Catchment Abstraction Management Strategies (CAMS), implementation of the Habitats Directive and the Water Framework Directive (WFD), and routine abstraction licensing and discharge consenting, all require the Environment Agency to estimate water resources potential for every catchment in England and Wales. This necessitates the estimation of river flows at both gauged and ungauged locations. For some of these activities, such as the setting of discharge consents and licensing of small abstractions, it is sufficient to encapsulate this information using a statistical description of the flow regime such as the flow duration curve. However, for the remaining activities, the use of a time series of river flows is either required, or would be advantageous.

The development of a model of the rainfall–runoff process that can be applied without recourse to calibration data, and hence applied within ungauged catchments to predict time series of river flows, has been the subject of research since the late 1960s. The theory is that, if the structural description is correct, the parameters of the model are more likely to be related to physical characteristics of the catchment that can be measured. Due to scaling issues and the many unknowns and sources of error, there are significant problems in retaining the physical basis of model parameters as models are scaled from the plot to the catchment scale. Hence, calibrated catchment-scale model parameters are commonly estimated in regionalisation studies through relationships (normally statistical) between parameters and datasets describing the physical and climatological structure of a catchment.

The problems in regionalising rainfall–runoff models are manifold; some key issues are summarised here. The extensive literature on calibration of rainfall–runoff models has shown that the parameters of a calibrated model are a function of the forcing data (including observed streamflow) and errors within these data, the model structure, the structure of the catchment being modelled and the calibration scheme (including the selection of objective functions). Hence model parameters may have little physical relevance. The catchment characteristics that can be used as predictors are nearly always extrapolated from point measurements, are inherently erroneous and may not be independent from one another. Furthermore, dependencies are rarely amenable to quantification through the use of classical statistical techniques such as covariance analysis.

Climate can play an important role in model descriptions of hydrological response, primarily due to limitations of models in describing the full complexities of the rainfall–runoff process. Hence, characteristics related to climate can commonly be useful predictors for model

parameters. This is not an attractive result as climate variables also form the input to the model and hence are not independent.

In the majority of previous studies, and certainly in the UK context, the identification of the fundamental limiting issues has been confounded by relatively small catchment sample sizes. Within the Modelling Historical Flow Sequences Project (Young *et al.*, 2000; Young, 2006), this was addressed through the exploration of the regionalisation problem over a large set of catchments and with two contrasting approaches to relating model parameters to catchment characteristics.

In the context of generalised rainfall–runoff modelling, the CERF project builds on previous work (Young *et al.*, 2000), by formally linking the structure of the designed rainfall–runoff model to the physical properties of a catchment through the development of a spatially structured model.

A primary question identified by the previous research, and the one that underpins CERF, was whether the approach of a priori calibration of lumped catchment models with the subsequent regionalisation of model parameters is the optimal approach to the problem of prediction in ungauged catchments. Undoubtedly part of the simulation error is introduced by the use of lumped models; nonetheless the advantage of lumped models over semi-distributed and fully distributed models is that lumped models have relatively few parameters to be identified from streamflow. However, the fundamental issue is whether the a priori choice of a lumped model structure is appropriate. In the approach adopted for CERF the detail of the model structure is determined by the physical characteristics of the catchment; this approach necessitates a semi-distributed model structure.

The CERF model was developed to simulate the time evolution of water resources, as represented by streamflow within a catchment. The temporal resolution for both managing and regulating water use is commonly between a week and a month. For assessing the yield of some storage schemes it is thus unnecessary to consider streamflow at a finer temporal resolution than this; for example, the estimation of natural inflows into an impounding reservoir. However, for run-of-river water use schemes, and cases where the impact of a scheme on aquatic flora and fauna is to be assessed, it is essential to consider streamflow at a finer temporal resolution. The commonly used resolution is a calendar day. This finer resolution is important for assessing the frequency of failure of a scheme (e.g. a direct abstraction), the ecological impacts of flow derogation and, in the case of discharges, the impact of flow on water quality and hence the flora and fauna of the stream.

In estimating the time series of natural streamflows at a site, it is not necessary to exactly replicate all aspects of the true natural time series. Within this project, the requirement was to simulate the most important facets of the flow regime. These include:

- an acceptable simulation of mean flow – conservation of mass;
- the behaviour of streamflow recession;
- the correct representation of seasonal patterns within the flow regime;
- the correct streamflow response to precipitation and the dependencies of that response on antecedent catchment conditions.

With regard to the last point, it is not important to simulate accurately individual high flow events. The only restrictions on the modelling of high flows are that mass should be conserved over a longer time period and the observed sequencing of high flow events should be replicated. In the context of run-of-river schemes, the high flows are not a resource that can be readily utilised due to problems with debris, high concentrations of suspended solids and other quality problems. The uncertainty in our knowledge of the cause–effect links between flow and habitat for aquatic species means that current

predictive methods for assessing the ecological impacts of high flows are not sensitive to the absolute magnitude of the flows.

As hydro-ecological and resource utilisation impact studies become increasingly important, the ability to estimate flow time series will become increasingly valuable. Similarly, the need to demonstrate, address and evaluate the potential impacts of climate and land-use change, will necessitate the use of deterministic time-series modelling tools such as those that have been developed within this study. For these reasons, the CERF model structure has been explicitly designed to allow the representation of both potential land-use change and climate change within ungauged catchments. The model has been calibrated and evaluated over a wide range (both spatial and temporal) of climatic conditions. There is no assumption of climatic stationarity in the input data and thus the modelling framework is appropriate for evaluating the consequences of climatic change. The model has been calibrated using recent land use and it could be argued that in some catchments land use may have changed significantly. However, the vegetation classes used within the model are relatively broad, based on water use rather than, for example, differences in cropping patterns, and are thus relatively insensitive to changes in agricultural crop type. Conversely, this does mean that the modelling framework can only be used to evaluate the consequences of major land-use changes associated with moving between vegetation classes (e.g. from arable to pasture or trees).

3 Summary of the CERF model

The CERF model is a rainfall–runoff model. The inputs into the model are time series of precipitation and potential evaporation demand and the output is a time series of simulated river flows. Traditional rainfall–runoff models have a set model structure; that is, the hydrological processes governing the translation of precipitation to streamflow are represented by discrete components with equations used to represent the hydrological processes within the components. This basic structure will not change between catchment applications and a traditional model is fitted within a catchment through the choice of parameters for the equations.

Within CERF the structure of the model does change between catchments. The basic model building block within CERF is a Hydrological Response Unit (HRU) which is applied to all parts of a catchment that are thought to 'behave' in a hydrologically similar manner. The rainfall–runoff process is broken down into three components: a loss module, a quick flow routing module and a slow flow routing module. Each module has a set of hydrological response units. Within CERF an HRU is defined by the catchment descriptors. The number of HRU used within a particular catchment depends on the complexity of the catchment. So, for example, a small catchment with very similar soils, geology and vegetation will have only a few HRU, while a large, diverse catchment will have many. The model parameters for each individual HRU are fixed and hence do not vary between catchments. The overall modelling strategy is presented as a flow diagram within Figure 3.1.

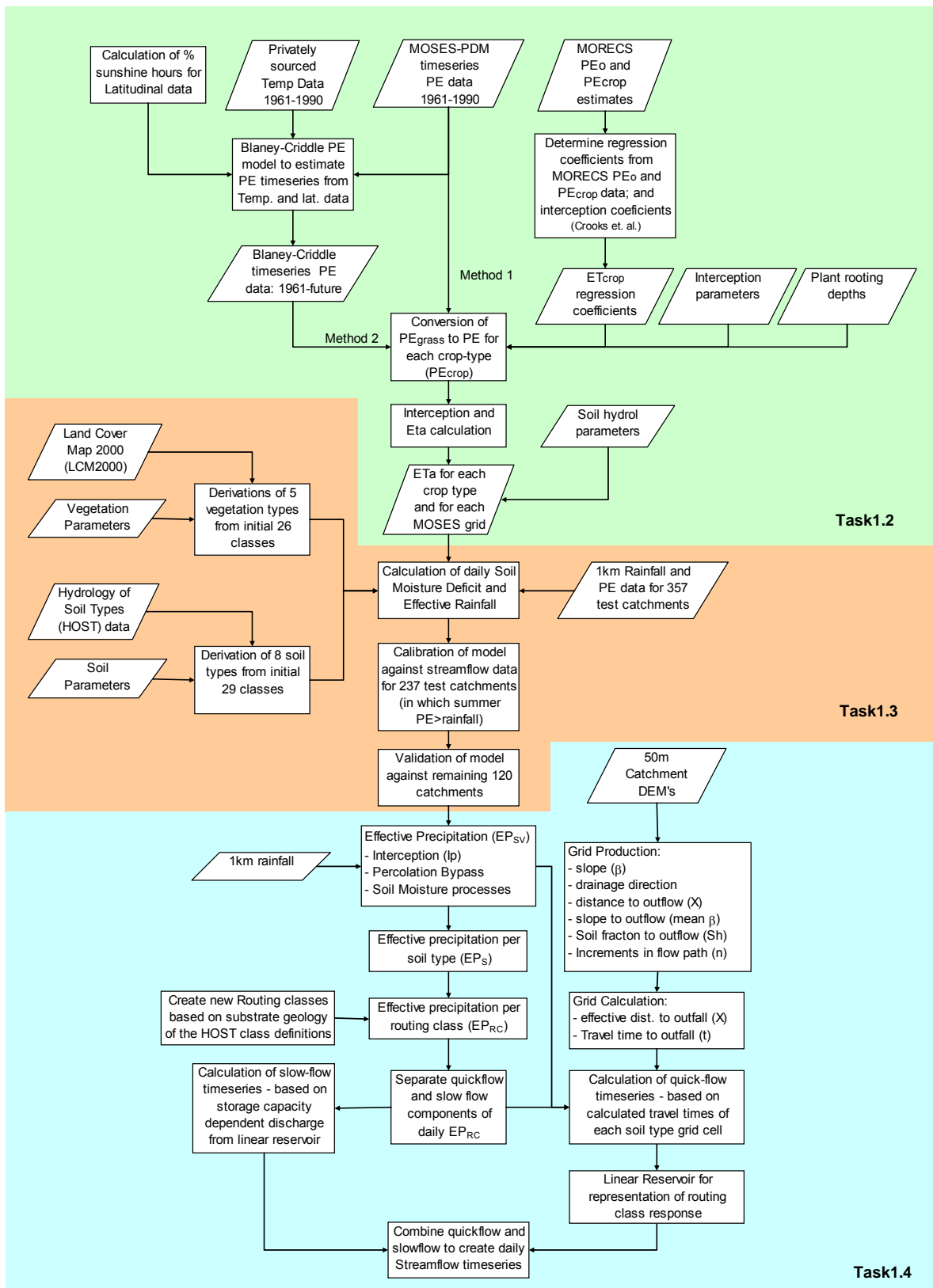


Figure 3.1 Flow diagram illustrating the three main stages of the CERF modelling process: estimation of evapotranspiration, soil moisture modelling scheme and semi-distributed routing model

The catchment descriptors used within the model structure are a hydrologically referenced 50-m resolution, digital terrain model (DTM) (Morris and Heerdegen, 1988); the Hydrology of Soil Types (HOST) 29 class hydrological response classification of soils across the UK (Boorman *et al.*, 1995) and a classification of land cover based on five broad vegetation classes (deciduous, coniferous, arable, grassland and upland) derived from the CEH Land Cover Map 2000 classification system mapped at a 50 m resolution (Smith *et al.*, 2001).

The model structure is based around two sub-model components: the loss module that generates hydrologically effective precipitation (EP) and the routing module that subsequently routes the EP to the catchment outlet. The basic model structure for the loss module is a hydrological response unit consisting of an interception sub-module and a treatment of transpiration losses based on the FAO56 soil moisture accounting procedures for determining crop water requirements (Allen *et al.*, 1998). The interception model was based on that regionalised for inclusion within a rainfall-runoff model by Young (2006). The model has one parameter: the maximum depth of water that can be held by the vegetation, γ . The conceptual structure of the FAO transpiration module is presented within Task 1.3 and summarised in Figure 3.2. The module describes vegetation as a function of maximum root depth, Z_r , and 'moisture depletion fraction', p , for a range of vegetation and soil types. The total available water (TAW), the amount of water available to plants after a soil has drained to its field capacity, is defined as the product of the difference between field capacity (FC) and wilting point (WP) (properties of the soil class) and Z_r . Plants freely transpires until soil moisture deficits (SMDs) exceed the threshold defined by $p \cdot Z_r$; beyond this threshold the plants become increasingly stressed and evaporation reduces below the potential rate in proportion to the depth of threshold exceedence. EP is generated by the module when the SMD within the module is zero.

HRUs were defined by combining the HOST classes and reduced land-cover classes to yield a potential 140 combinations of HOST and land-cover classes plus an open water class. The number of actual combinations was significantly less as some land-cover/soil class combinations do not occur in practice. At the catchment level the individual cells within the HRUs represented within the catchment are amalgamated to form HRUs with a fractional extent that is not necessarily contiguous within the catchment. The response of each HRU is controlled by the vegetation parameters of γ , Z_r and p and the FC and WP parameters for the soil class. FC and WP parameters were defined for each soil class based on extracting the average percentages of sand, silts and clays within each HOST class from the UK National Soil Resource Institute's SEISMIC dataset. This process leaves Z_r and p , for each vegetation class, as the free parameters for the HRUs within the loss module which equates to ten parameters in total.

The output from the loss module within a catchment is an EP time series for each 1 km cell within the catchment. The routing module routes these effective precipitation time series to the catchment outlet via a semi-distributed routing scheme. Within the UK the dominant influence on the routing of water through the land surface is hydrogeology and its impact on soils and topography.

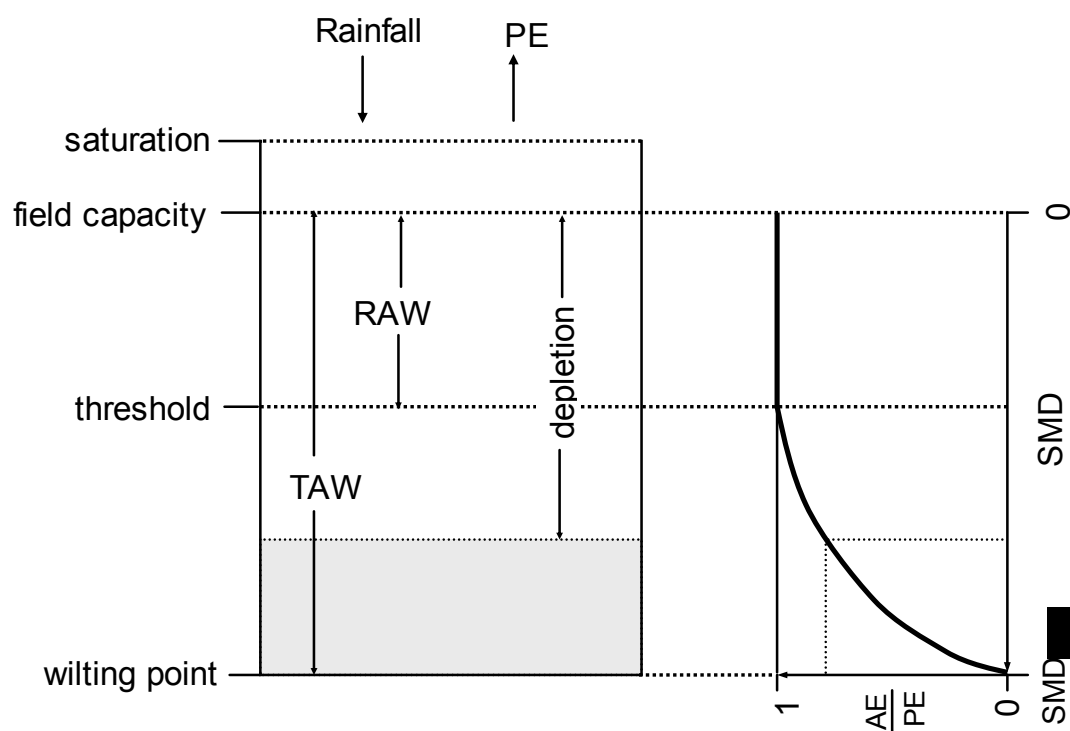


Figure 3.2 Soil moisture accounting structure

In the absence of an appropriate, complete high resolution digital hydrogeological classification of the UK the 29 HOST classes were amalgamated into 11 hydrogeological routing HRUs based on substrate geology. The EP time series for each cell enters the routing HRU corresponding to the HOST class of the cell. The routing HRU consists of a probability distributed storage representing free water in the soil column coupled to quick and slow flow routing pathways. Direct runoff from the probability distributed storage passes through the quick flow routing path while drainage from the store passes through the slow flow routing path. The quick flow routing path is based on topographic routing of this component of the effective precipitation to the catchment outlet while lumped, linear reservoirs are utilised within the slow flow pathway.

The probability distributed storage is assumed to be uniformly distributed with a maximum storage depth of 75 mm (determined by preliminary individual catchment model applications), drainage takes place from the base of the store and is proportional to the depth of water held in storage. The constant of proportionality, K_g , is a free parameter. The runoff from the store is routed through a topographically defined routing path while the drainage is routed through a linear reservoir, with a time constant K_b , representing the base flow for the routing class.

The previous regionalisation work on lumped models, in which quick flow routing was via a linear reservoir, demonstrated a strong dependency of the reservoir time constant on both catchment size and soil storage. For this study the quick flow routing within the routing HRUs was subdivided into a topographic component and a component representing transient soil storage along the routing path. The topographic routing of the quick flow from the individual cells within a routing HRU to the catchment outlet was defined based upon the flow path defined from the DTM and the cell-level topographic gradients along the path. A soil transfer velocity is included within the routing path, which conceptually is linked to bulk, lateral hydraulic conductivity. The total topographically routed quick flow for the routing HRU is calculated as the sum of the EP time series for the constituent cells lagged by the corresponding cell travel times.

The resultant summed time series is then passed through a linear reservoir to represent the transient storage along the flow path lengths.

A comparison of transposed flow data with the predictive performance of the CERF model illustrated that in the majority of cases, where the analogue and ungauged catchments are un-nested, the use of the CERF rainfall–runoff model will generally give better results than the use of data from the analogue site. If the analogue catchment is nested (either upstream or downstream) with the ungauged catchment (and relatively close), the transposing of gauged flow data generally yields a better time-series fit than that obtained with the CERF rainfall–runoff model.

A comparison of the predictive capacity of the CERF model with that of Low Flows 2000 for the prediction of average annual runoff demonstrated that the performance in predicting average annual runoff is comparable for both methods with the biggest differences occurring in dry (rainfall less than 800 mm per year) catchments. Low Flows 2000 predictions of Q_{95} were demonstrated to be superior to those from the CERF model. This is not a surprising result as the CERF model is structured and calibrated for whole hydrograph prediction (rather than estimation of flow duration curves) while Low Flows 2000 is specifically structured and calibrated for the prediction of average annual runoff and flow duration statistics.

4 Task 1.1: Estimation of precipitation inputs

The overall objective of this task was to develop 1-km resolution grids of daily rainfall for the period from 1961 to 2001 for the UK and to provide a methodology for updating the rainfall grids. The spatial interpolation algorithm used was that adapted for use within earlier work (Young *et al.*, 2000) based on a triangular planes methodology. The activities defined under this task were as follows:

- To evaluate the sensitivity of the accuracy of the spatial interpolation algorithm for generating rainfall grids to the density of the rain gauge data.
- To develop daily grids of 1 km² rainfall based on the appropriate selection of rain gauges and methodological approach and for the period 1961 to 2001.
- To develop full documentation of the methodologies so that they can be used to estimate rainfall grids using rainfall data prior to 1961 if required.
- To review the literature on rain gauge under-catch and comment on the potential implications on the accuracy of the rainfall grids.

A copy of the Meteorological Office daily rainfall library (containing over 6000 recording gauges within Great Britain), held by CEH, was used as the source dataset for deriving the daily rainfall grids. The derivation of these rainfall grid time series was extremely computationally intensive. The triangular planes methodology used, was coded within FORTRAN within a parallel processing batch format in order to facilitate the use of the CEH GRID parallel processing network.

Assuming an unbiased distribution of recording gauges and considering a range of catchment sizes, the analysis demonstrated that accuracy (likelihood of systematic error) of catchment average rainfall time series is insensitive to rain gauge density, but precision is sensitive to gauge density as the correct representation of individual precipitation events within the time series is obviously sensitive to the density of measurement.

Systematic errors in rainfall (i.e. bias) can lead to bias in the values of calibrated model parameters as the choice of parameters that give the best model fit will compensate for the bias in the input data. While the overall quality of simulation is sensitive to the precision of the rainfall time series, the choice of model parameters is less so. Thus, unless the distribution of rain gauges is biased the parameterisation of the CERF model should not be that sensitive to the inclusion of catchments with a sparse rain gauge network within the calibration dataset.

The importance of rain gauge under-catch was reviewed based upon peer-reviewed evidence and experience gained with the development of Low Flows 2000. This evidence suggested that rain gauge under-catch may contribute to a small tendency to under-predict average annual runoff. However, this is not demonstrated within lower rainfall catchments in which rain gauge under-catch is not a critical factor, and where such systematic error can be compensated for by the soil moisture accounting model (without affecting the accuracy of parameter estimates).

The following deliverables were produced in this task:

- 1-km resolution daily rainfall grids for the UK.
- Model code for the interpolation of 1-km resolution daily rainfall grids from the appropriate selection of rain gauges for the period 1961 to 2001.
- Full documentation of the methodologies so that they can be used to estimate rainfall grids using rainfall data outside the period 1961 to 2001, if required.

5 Task 1.2: Model scheme for representing evapotranspiration processes

The overall objective of this task was to develop a consistent methodology for the estimation of potential evaporation (PE) from different land-cover types. The individual activities within this task consisted of the following:

- The selection of a set of UK climatologically representative case studies for evaluating the sensitivity of model outputs to PE representation.
- A review and comparison of the Met Office Rainfall and Evaporation Calculation System (MORECS), and the Met Office Surface Exchange Scheme with the CEH Probability Distributed Moisture scheme (MOSES-PDM).
- The review and development of an appropriate land-cover classification using the national CEH land-cover datasets.
- The calibration of models against observed data using high quality PE data based on the most appropriate source of PE data.
- An evaluation through case studies of the sensitivity of the streamflow simulations, soil moisture and actual evaporation state variables within a representative rainfall–runoff model to different estimation procedures for PE.
- The development of gridded datasets of PE data at an appropriate spatial and temporal resolution for use within the project.

In the first instance, the MOSES-PDM reference daily PE data were obtained for the short-grass reference crop for the period 1961–2001. After communication with the National Groundwater Recharge Assessment project (W6-089), PE data for the five other hydrologically relevant land-cover classifications derived from the CEH Land Cover 2000 dataset were determined (coniferous forest, deciduous forest, winter cereal, summer cereal, and upland vegetation types). Conversion of the MOSES-PDM PE data into PE estimates, for these vegetation types, was achieved using empirical regression equations based on MORECS PE estimates for each land-cover type at nine contrasting locations within the UK using data for the 1961–1990 standard period.

As an alternative to the use of MOSES-PDM data, a temperature-based PE estimation method based on the Blaney-Criddle model was also developed over the same standard period. The modified Blaney-Criddle model was developed so that users with a lack of access to MOSES-PDM data will still be able to employ the CERF model for streamflow time-series estimation. Figure 5.1 illustrates the input data requirements for both modelling approaches and subsequent calculation of actual evapotranspiration for different crop types.

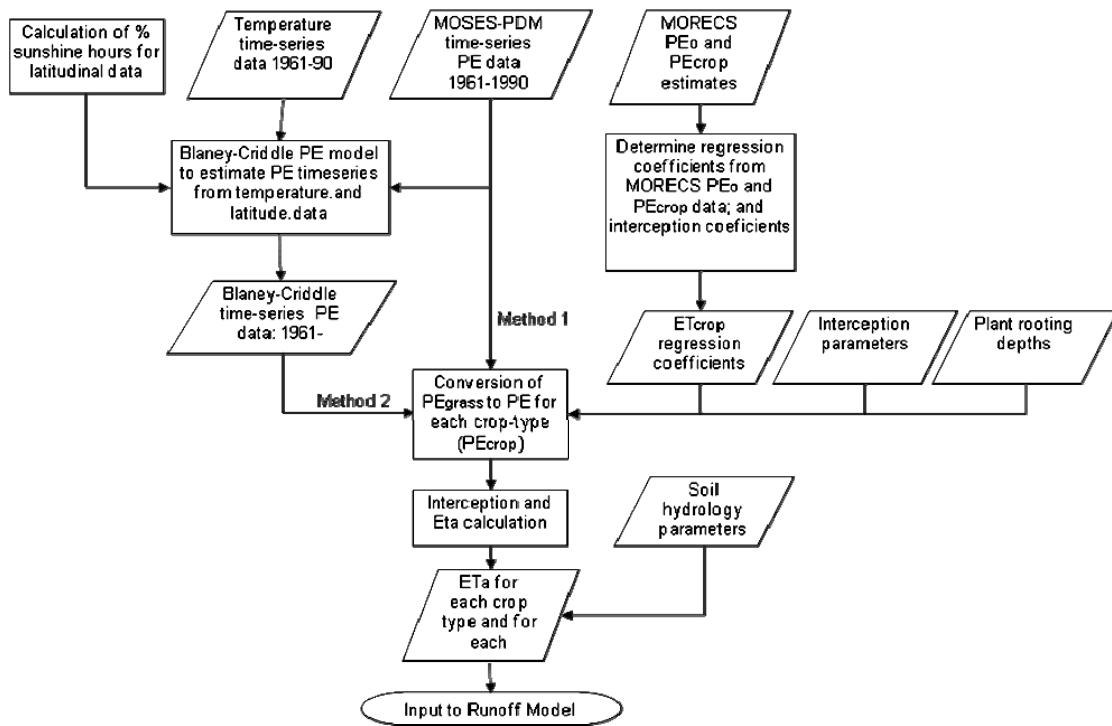


Figure 5.1 Calculation of actual evapotranspiration for different crop types from PE estimates and land-cover parameters.

An evaluation of rainfall–runoff model sensitivity to PE data produced by the different estimation procedures described above was conducted for a selection of climatologically representative case studies for the UK. The analysis revealed that the simulated river flows were not sensitive to the PE estimation procedure used. Consequently, although MOSES-PDM data were used for the development of the CERF methodologies, the CERF model may be used in conjunction with PE estimates derived using the Blaney-Criddle model and possibly other schemes without any significant loss of predictive performance.

The MOSES-PDM data used for the project were hindcast data for the period 1961 to 2001 based on MORECS forcing data and thus were constrained to a 40 km grid. This obviously can result in boundary discontinuities within smaller catchments. To address this point when deriving catchment-averaged values of PE from the 40 km resolution PE data temporary 1-km resolution grids were constructed by interpolating from the four nearest 40 km grid centroids using a weighted average of the four values. The weights used were the inverse distances of the 40 km centroids from the centroid of the target 1 km cell.

The following products were delivered from this task:

- Spreadsheets of daily PE for each Met Office defined 40 km² PE grid square in the UK (as derived from Met Office Access database) for the period 1961–2001.
- Model code for the interpolation of catchment averaged daily PE from 40 km² PE.
- Full documentation of the above methodologies so they can be used to independently estimate PE for any catchment or time period if required.

6 Task 1.3: HRU model scheme for representing rainfall interception and soil moisture

This component of the model includes the calculation of daily interception, actual evaporation (AE), soil moisture deficit (SMD), and effective precipitation (EP), from daily rainfall and potential evaporation (PE) time series. As discussed within Section 3 the model structure for each component is based on the definition of Hydrological Response Units (HRUs) consisting of an interception model coupled to a soil moisture accounting procedure based on the FAO56 procedure (Allen *et al.*, 1998). The basic building block for the HRU is a combined soil-vegetation classification. The activities within this task consisted of the following:

- Development and testing of an appropriate HRU soil moisture model structure (as defined by the above catchment characteristics).
- Development of a combined classification system for land-cover and soil types.
- Development of an appropriate multi-objective function calibration scheme and the simultaneous calibration of the HRU structure across many catchments.

A simplified land-cover classification system was developed from the CEH Land Cover Map (LCM2000). This coverage consisted of grassland, arable, upland, urban, coniferous forest and broadleaf forest. The distribution of 29 soil classes across the UK was obtained from the CEH Hydrology of Soil Types (HOST) database.

A one-parameter interception storage model for deciduous and coniferous land-cover types previously regionalised by Young (2006) was further developed and recalibrated at a number of sites across the UK using field data on measured interception losses and from previous interception studies and the spatially interpolated rainfall grids produced in Task 1.1.

Precipitation that is not intercepted enters the soil moisture store. Evaporation takes place from this store at a rate which is a function of the PE demand within the time step, the antecedent soil moisture deficit, land-cover type, vegetation rooting characteristics and soil properties (field capacity and wilting point) using a modified form of the FAO56 accounting procedure. A mass balance is calculated across the store within a time step and effective precipitation is generated from the store when the storage deficit is zero and the store spills in response to precipitation.

The HRU structure is defined entirely by the combination of vegetation class and soil class within a 1-km resolution cell. Cells with the same combination of vegetation and soil classes are grouped together to form the HRU for that soil-vegetation class. As discussed within Section 3 each HRU has four parameters: the maximum depth of water that can be held by the vegetation, γ , the maximum root depth, Z_r , the 'moisture depletion fraction', p , and the field capacity (FC) and wilting point (WP) for the soil class. As discussed the interception parameter, γ , was calibrated against field data while values of FC and WP were taken from the pedological transfer functions used to ascribe values of these parameters to soil classes within HOST. The values of the remaining vegetation parameters of maximum rooting depth and moisture depletion factor were optimised across 177 UK catchments for which reliable streamflow records were available and within which significant summer soil moisture deficits were known to occur.

As each catchment contained a different configuration of soil and vegetation classes multiple model–parameter combinations, were optimised simultaneously. A linear nested parameter search algorithm was used to identify the optimal parameter set for each HRU using model forcing data from the period 1987 to 2001.

Model performance was evaluated over the available data within the period from 1960 to 1986 for the calibration set of catchments and evaluated over an independent set of 157 catchments in which summer rainfall exceeded summer PE. The model was found to perform equally well in both instances.

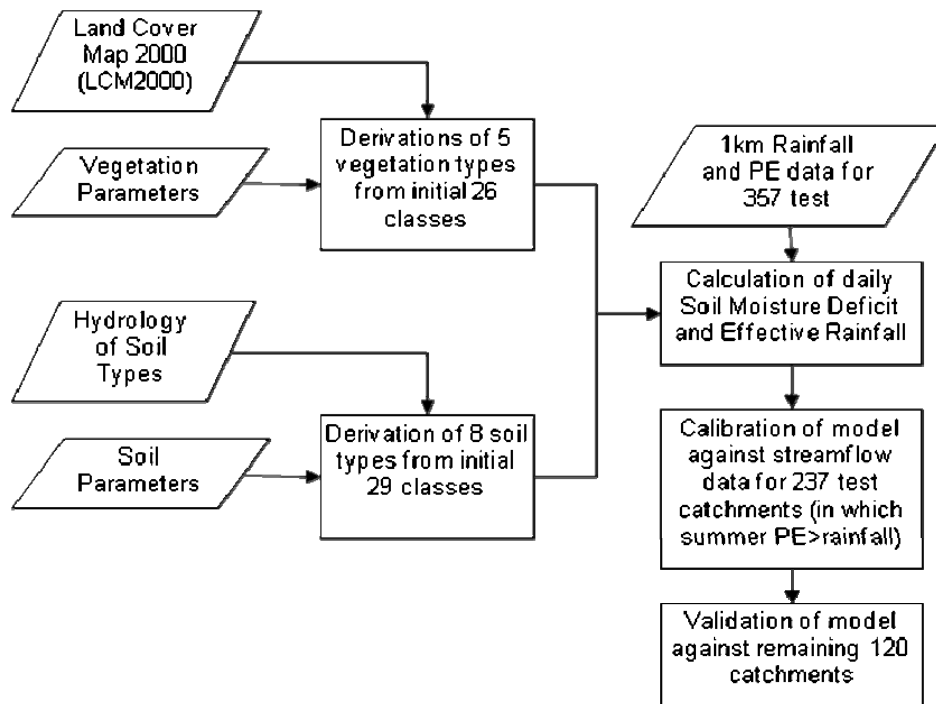


Figure 6.1 Use of LCM2000 and HOST data, and subsequent calibration and validation of the soil moisture model using data from 357 identified test catchments.

The following products were provided from this task:

- Model code for the interception and soil moisture components of the model.
- Model calibrations for all catchments within the dataset.
- Description of methods for estimating parameters directly from catchment characteristics.
- Test catchment datasets including catchment boundaries, soil types and vegetation coverages.

7 Task 1.4: Model scheme for routing effective precipitation to streamflows

The overall objective of Task 1.4 was to develop a generalised model scheme for routing effective precipitation to streamflows with an evaluation of the prediction uncertainty. In developing the final model scheme both lumped and semi-distributed approaches were evaluated. The activities associated with this task comprised:

- Development and evaluation of the lumped routing scheme, lumped at the sub-catchment level, based on a simple second-order linear reservoir routing scheme.
- Development and evaluation of the semi-distributed scheme.

The level of performance of both the semi-distributed and the lumped routing scheme were assessed and compared. The predictive performances of both routing schemes were very comparable; the lumped scheme gave marginally greater performance in terms of high flow prediction while the semi-distributed scheme is more effective in simulating the low flows.

The development and evaluation of these activities are reported in detail within the corresponding technical report. The semi-distributed scheme was selected for the final CERF model as it is also transparent and thus, if it is implemented fully, users will be able to explore how runoff is generated from different parts of the catchment. With appropriate training, this transparency should lead to greater confidence in the subsequent use of the model within ungauged catchments. The use of the semi-distributed scheme also opens up the potential for the wider use of the CERF model within the Environment Agency for underpinning contaminant transport models.

As discussed in Section 3, the semi-distributed routing scheme is based on the topographic routing of quick flow and a simple representation of slow flow components based on soil characteristics (Figure 7.1).

In the absence of an appropriate resolution digital hydrogeological classification of the UK, the 29 HOST classes were amalgamated into 11 hydrogeological routing HRUs based on substrate geology. The EP time series for each cell enters the routing HRU corresponding to the HOST class of the cell. The routing HRU consists of a probability distributed storage representing free water in the soil column coupled to quick and slow flow routing components. The probability distributed storage is assumed to be uniformly distributed with a maximum storage depth of 75 mm (determined by preliminary individual catchment model applications), and drainage takes place from the base of the store and is proportional to the depth of water held in storage. The constant of proportionality, K_g , is a free parameter. The runoff from the store is routed through a topographically defined routing path while the drainage is routed through a linear reservoir, with a time constant K_b , representing the baseflow for the routing class.

The previous regionalisation work on lumped models (Young, 2006), in which quick flow routing was via a linear reservoir demonstrated a strong dependency of the reservoir time constant on both catchment size and soil storage. For this study the quick flow routing within the routing HRUs was subdivided into a topographic component and a component representing transient soil storage along the routing path. The topographic routing of the quick flow from the individual cells within a routing HRU to the catchment outlet was defined based upon the flow path defined from the CEH digital terrain model (DTM) and the cell-level topographic gradients along the path. A soil transfer velocity is included within the

routing path, which conceptually is linked to bulk, lateral hydraulic conductivity. The total topographically routed quick flow for the routing HRU is calculated as the sum of the EP time series for the constituent cells lagged by the corresponding cell travel times.

The resultant summed time series is then passed through a linear reservoir to represent the transient storage along the flow path lengths.

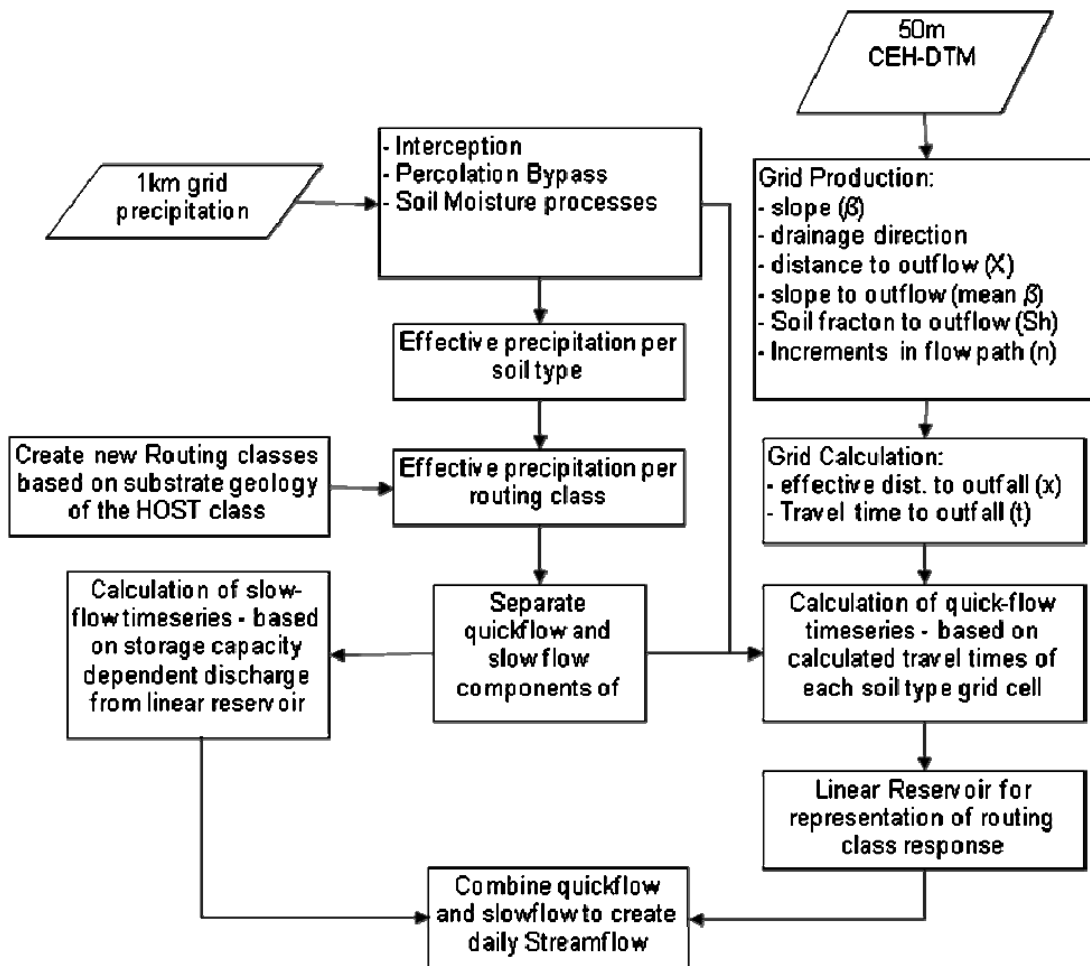


Figure 7.1 Flowchart of semi-distributed structure approach

The objective functions used to measure the overall model performance during calibration and evaluation were the:

- overall model BIAS (error in predicting mean flow);
- Nash-Sutcliffe efficiency criterion (NSE) (Nash and Sutcliffe, 1970);
- mean sum of squared differences between observed and simulated flows over the observed flows located within the lowest third of the observed flow distribution (LF_OBJ);
- simulation bias error at the observed Q_{95} flow (the flow that is equalled or exceeded for 95% of the time), BEQ95.

BIAS measures the ability of the model to close an overall water balance, while the use of the Nash-Sutcliffe Efficiency Criterion reflects the requirement to model the general variation in the time series accurately. The LF_OBJ objective function reflects the

requirement to model recession periods (i.e. periods of low flow) precisely and the BEQ95 statistic measures whether or not there is a systematic simulation error at low flows.

The performance of the model within the development catchments is summarised within Table 7.1 over both the calibration and evaluation periods and in Table 7.2 within the corresponding evaluation catchments.

Table 7.1 Performance comparison between the semi-distributed and the lumped routing scheme over the calibration period within the common catchments used for calibration (138 catchments)

	Calibration period			Evaluation period		
	LF_OBJ	NSE	BEQ95	LF_OBJ	NSE	BEQ95
68%c.i. l.l.	35	0.60	-14	33	0.54	-19
median	60	0.73	59	57	0.69	42
68%c.i.u.l.	127	0.80	171	114	0.77	135

These tables demonstrate that the quality of model fit is extremely stable both between calibration and evaluation periods and between the development and independent catchment sets. Furthermore, the quality of fit is good within relatively low values of LF_OBJ and BEQ95 and high values of NSE.

Table 7.2 Performance comparison between the semi-distributed and the lumped routing scheme over the calibration period within a set of independent evaluation catchments (70 catchments)

	Calibration period			Evaluation period		
	LF_OBJ	NSE	BEQ95	LF_OBJ	NSE	BEQ95
68%c.i. l.l.	25	0.58	-19	28	0.51	-14
median	35	0.71	15	37	0.67	17
68%c.i. u.l.	57	0.78	54	55	0.77	63

Example hydrographs for three catchments for a contrasting range of geology and climate are presented within Figure 7.2. The locations of the catchments is presented within Figure 7.3.

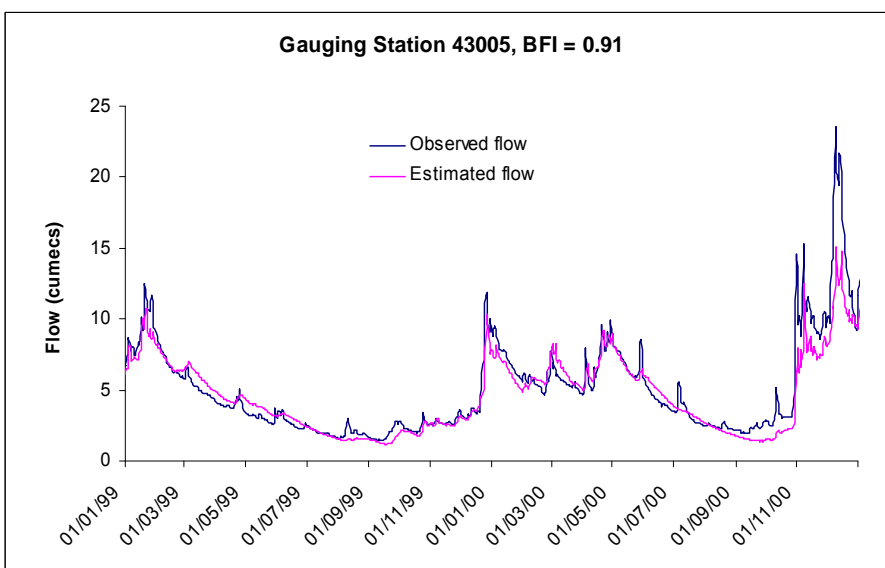
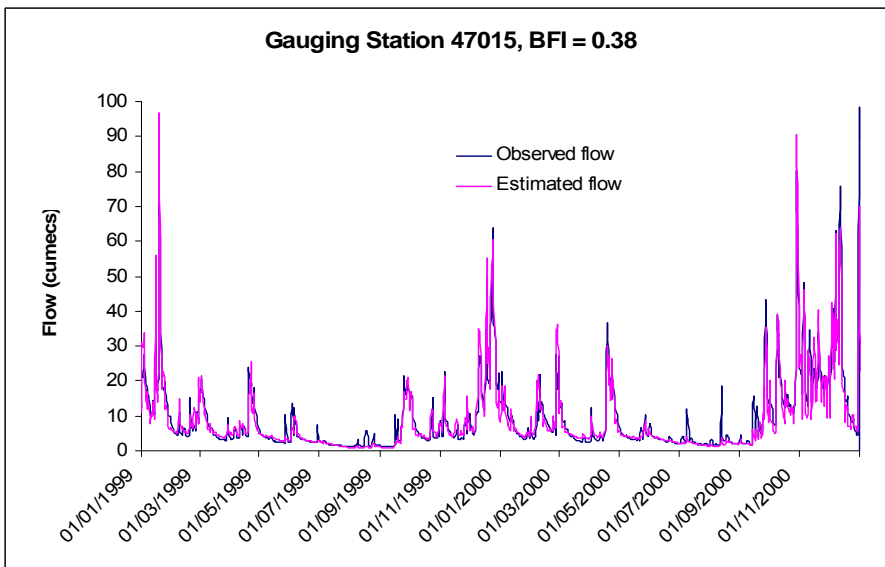
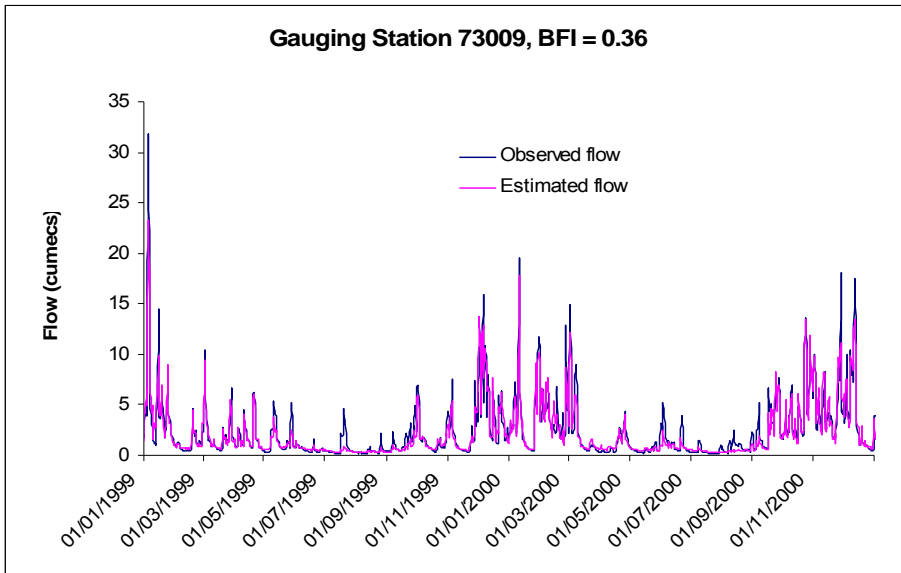


Figure 7.2 Simulated and observed hydrographs for the example catchments



Figure 7.3 Location of example catchments

The following products were delivered through this task:

- Model code for estimation of streamflow time series.
- Model code for statistical analysis of modelled streamflow time series.
- Test catchment observed and modelled time series.
- An evaluation of the dependency of the predictive performance of the CERF on the routing option, lumped or semi-distributed, used.

8 Task 2.1: An evaluation of the sources of predictive uncertainty within the CERF model

In application the quality of the simulation of actual streamflow provided by a rainfall–runoff model will be influenced by a number of factors including:

- incorrect assumption of a closed catchment water balance (a fundamental assumption within most rainfall–runoff models including CERF);
- errors in forcing data, which for the CERF model consist of input precipitation, potential evaporation and, in the context of the model development, streamflow data;
- error in the estimation of contributing catchment area and whether a closed catchment water balance is an appropriate assumption;
- incorrect model parameterisation and/or model structure.

In the traditional rainfall–runoff model application where the model is applied within a catchment with measured streamflow data, and is calibrated against the streamflow data, predictive performance is measured through comparison of the simulated and measured data over the calibration period. Usually, the stability of the model simulation is then assessed by comparing the model performance against measured streamflow from outside the calibration period.

The CERF model has been developed for use within catchments for which there are no measured flow data and hence a guide to the likely performance of the model within a catchment can only be gained by comparing the performance of the model in similar catchment types with a similar quality of input data. As discussed within the previous section, the performance of the CERF model is generally good across a wide range of catchments. However, as with all models, and particularly regionalised models, there are examples of catchments in which the model performance is not so good.

The overarching objective of this task was to investigate and evaluate the sources of error that may arise within the application of the CERF model in an ungauged catchment.

To facilitate this analysis the CERF model was applied to the much wider pool of 647 gauged catchments used within the Low Flows 2000 station pool. It should be noted that the Low Flows 2000 assessment criterion for artificial influence is that the influenced ratio of Q_{95} to mean flow should not be significantly different from the natural ratio. It is the case that this station pool contains catchments that are significantly influenced. The quality of the simulation was evaluated, and in cases where model performance was poor the contribution of the range of possible sources of error listed above were considered.

The performance of the model was quantified using the three objective functions: overall bias (BIAS), Nash-Sutcliffe efficiency criterion (NSE), and bias at Q_{95} (BEQ95) (see previous section). Good simulations were defined as those meeting the following set of criteria:

- The BIAS had to be less than $\pm 30\%$ corresponding to the 95th percentile confidence interval for the Low Flows 2000 prediction of mean flow.

- The NSE value had to be greater than 0.5, indicating that the simulation is explaining more than 50% of the variation in the observed data.
- The BEQ95 value had to lie between -60% and +120%. This range was selected due to the asymmetric nature of this bias error which has a maximum low limit of -100% representing a theoretical simulation of zero flow at Q_{95} but no theoretical upper limit. It should also be noted that there is a dependency on catchment type; in flashy, low Q_{95} catchments the BEQ95 error will be higher for a given absolute error than in less flashy catchments.

The simulations within catchments that failed at least one of these criteria were examined. The catchments were then classified according to the level of performance and, if poor, whether the causes of failure were identifiable. The results from this classification are presented in Table 8.1 and are examined in more detail within the Technical Report 2.1.

Table 8.1 A classification of catchments depending on their performance level

Class	Good simulation	Potential source of error	Number
AA	Y		345
B1	N	Inputs – unreliable catchment area	20
B2	N	Inputs – unreliable precipitation	32
B3	N	Inputs – anomalous observed flow	46
C	N	Model regionalisation – chalk	27
D1	N	Model structure – ephemerality	62
D2	N	Model structure – other karstic features	8
U	N	Unexplained	107

The model produces good quality simulations (i.e. passes all three criteria) within 345 of the catchments. This dataset includes the majority of the catchments used for calibration and initial testing of the model (see previous section). It should be easier to close a water balance in the development catchments as they have the best hydrometric data and are the most natural catchments.. Many of the remaining catchments will have lower quality hydrometric data and will be subject to greater disturbance. We return to this point later. It should also be recognised that while the model may violate the limits set for one objective function it may still return very good simulation values for the others and the failure of one might not necessarily compromise the subsequent use of the simulated time series. It is therefore an incorrect conclusion that the CERF model does not work well within the remaining catchments and hence the quality of simulation should be evaluated in the context of what is required from the simulation for a particular application (see Task 2.3).

The catchments in Class B1 were all catchments in which the topographic area identified using the DTM differed significantly (by more than 10%) from the contributing area reported by the measuring authority. This highlights the importance of ensuring the best possible estimate of contributing area is used in the application.

Erroneous precipitation time series were also found to be a significant source of error in 32 catchments. Analysis of the rain gauge density in these catchments showed that these

errors tend to occur where the average distance between rain gauges that contribute to the estimate of precipitation is large compared with the scale of the catchment. This is commonly the case in small catchments within upland areas.

Class B3 contains stations in which the National River Flow Archive station notes identify hydrometric issues, such as flood bypass or drowning of the structure, and also stations for which the gauged runoff is anomalous compared with that for similarly sized catchments within the region. These anomalies could potentially be related back to water balance errors (Class B1) but without specific evidence to support this they were retained in Class B3.

The problems observed in simulating the chalk catchments within Class C are common to all regionalised models and in many cases reflect the incorrect assumption of a closed catchment water balance resulting from gauge bypass or the presence of karst features such as spring lines. The catchments within Class C could have been reclassified within other classes but were retained as an individual class to highlight the problems of applying regionalised models within these systems. In practice, many of the problems in this class could be resolved to a great extent by a bespoke, basin-level local calibration of CERF against basin-level gauged data and then subsequent application of the locally calibrated model within ungauged catchments.

The model simulations within Class D generally failed on the simulation of low flows only, but within many of the catchments in the class the simulations would be perfectly acceptable for many water resources applications. The CERF model cannot predict zero flows and while the model may simulate very low flows these will invariably represent a large percentage of the observed flows at Q_{95} if the catchment is starting to dry out.

There are a large number of catchments in which the source of the simulation error could not be identified. One very important issue is the potential perturbation of the gauged records within the catchment dataset resulting from artificial influences. The overall performance of the CERF model is lower than that observed within the pool of stations used to develop the model (see previous section). The catchments selected for developing the model were essentially those catchments from the Low Flows 2000 gauging station pool judged to be both hydrometrically good and essentially natural. The wider pool used for this analysis includes a large number of stations for which the ratio of the Q_{95} flow to the mean flow is considered to be similar to the corresponding natural ratio but which may be significantly influenced in absolute or temporal terms. It is recommended that the analysis undertaken within this task is repeated, prior to implementation, using the revised station pool that, at the time of writing, is currently being developed by Wallingford HydroSolutions Limited (the licensed providers of Low Flows 2000).

The product from this task is the Technical Report 2.1 detailing the full analysis.

9 Task 2.2: Guidance for prediction method selection based on minimisation of uncertainty

The CERF model provides one approach to predicting time series of streamflow within ungauged catchments. More sophisticated tools, such as distributed groundwater models, may have been calibrated across the basin. In the absence of such a model the other viable alternative to CERF for the prediction of time series is to transpose flow data from a suitable analogue catchment. Potentially time series from the CERF model could be summarised for application as flow duration statistics. In this case these flow duration statistics could also be derived directly using the Low Flows 2000 software system. The activities associated with this task were:

- To compare the predictive capacity of the CERF model with transposed gauged flow data from an analogue catchment, and to provide guidance on whether to use CERF or the transposed record when suitable analogue records are available.
- To compare the predictive capacity of the CERF model with that of Low Flows 2000 for the prediction of average annual runoff and the Q_{95} flow duration statistic.

The transposition of data from nested catchment pairs (either upstream or downstream) meeting relatively stringent similarity measures generally results in very good simulations of the time-series structure of observed flow data at the target site. By contrast, the quality of simulations obtained by transposition of data from un-nested analogue catchment pairs is generally significantly worse.

The performance of the CERF model for target catchments was described relative to the performance of the transposition method used in both types of analogous catchment. The comparison of transposed flow data with the predictive performance of the CERF model illustrated that, in the majority of cases, if the analogue catchment is connected to the ungauged catchment (nested with less than factor 4 difference in contributing area), the transposing of gauged flow data generally yields a better time-series fit than that obtained through using the CERF model. In cases where the analogue and ungauged catchments are un-nested, the CERF rainfall–runoff model will generally give better results than the use of data from the analogue site.

The use of flow data from an un-nested analogue can result in very larger errors both from the perspective of replicating the time series and the closing of an adequate water balance. Conversely, the relatively good replication of the time series for the nested dataset is a consequence of the strong serial correlation between the flows observed at the target site and the analogue site.

The analysis relied on both the target and analogue catchments having relatively natural flow regimes. Ideally, a wider range of catchment pairs, including pairs that were similar hydrogeologically, but had significantly different catchment area, would have been used. However, as a consequence of the requirement for naturalness, selected pairs were inevitably biased towards the smaller headwater catchments. This is particularly the case for the selection of nested pairs and resulted in analogue pairs being very similar in scale. The consequence of this was that the selected analogues presented an extremely good set

of analogues, which will influence the analysis. In practice analogues of this quality may be difficult to locate in many applications.

The comparison of the predictive capacity of CERF with that of Low Flows 2000 for the prediction of average annual runoff demonstrated that the performance in predicting average annual runoff is comparable for both methods, with the biggest differences between observed and predicted occurring in dry catchments for both models. Low Flows 2000 predictions of Q_{95} were demonstrated to be superior to those from CERF. This is not a surprising result as CERF is structured and calibrated for whole hydrograph prediction while Low Flows 2000 is specifically structured and calibrated for the prediction of average annual runoff and flow duration statistics.

The following deliverables have been produced in this task:

- Guidance on the selection of the best method for estimating time series of flows in ungauged catchments in relation to minimisation of expected prediction uncertainty.

10 Task 2.3: Time-series indicators of operational utility

Rainfall–runoff models are calibrated using objective functions to evaluate different aspects of predictive performance. However, these objective functions are rarely related to the utility value of the simulated time series to potential users. The utility of the CERF model to the Environment Agency will be strongly influenced by the usefulness of the produced simulations in supporting Environment Agency business. It is therefore important to quantify simulations using measures that reflect their potential value to associated business processes.

This task developed a number of 'time-series indicators of operational utility' and explored the sensitivity of each of these indicators to different sources of error that might influence the CERF model. Observed streamflow time series were perturbed using mathematical functions to reflect different sources of error (rainfall input error, structural error in soil moisture model and parameterisation of the routing model). The quality of the perturbed time series were then evaluated using the same objective functions that were used within the CERF model calibration process.

The following time-series indicators used within the work were initially proposed by CEH, and subsequently discussed with the CERF technical board and representatives of the national hydrology and hydro-ecology teams within the Environment Agency:

1. FDC statistics – Q_{95} , Q_{70} , Q_{50} , Q_{20} , Q_{10} , Q_5 and Q_1 . These statistics provide information on whether the variance of predicted flow time series is replicated by the model and also serve as a useful comparison with the performance of Low Flows 2000.
2. Censored Nash-Sutcliffe Efficiency (NSE) – computed for flows observed below specified percentiles (Q_{20} , Q_{10} , Q_5 , Q_4 , Q_3 , Q_2 and Q_1). This was used to assess the time-series fit at flow ranges relevant to water resources.
3. R^2 of observed against predicted monthly mean flow time series – this is used to evaluate how accurately the model is replicating the underlying monthly flow signal and was calculated for summer and winter (seasonal) time series.
4. R^2 of observed against predicted monthly Coefficient of Variation (CoV); that is, the ratio of the standard deviation to the mean. This statistic was used to evaluate whether the model is preserving the variance of monthly data. The CoV for each month is the standard deviation of the flows within the month divided by the mean flow during that month. Again, this test will be repeated for summer and winter (seasonal) time series.
5. Predictive performance of the model in dry and wet years normalised to predictive performance over a range of years – used to assess model simulation of extreme years. This is particularly important if the model is to be used as a predictive tool for extreme flows (i.e. drought or flood forecasting).

The procedure was repeated for a number of case study catchments of contrasting geological and hydrological character. The sensitivity of objective function statistics and time-series indicators of operational utility to perturbed time series was found to depend on the hydrogeological and climatic characteristics of the catchment being modelled.

Catchment types, in which model predictions may be more sensitive are summarised below with respect to each type of introduced error.

Rainfall estimation errors

- The sensitivity of model results to biased and unbiased errors within input rainfall data will be most significant within low baseflow index (BFI) catchments, where streamflow is more responsive to rainfall variation.
- Model predictions within high BFI catchments, where the streamflow response is more damped, will be less sensitive (except during very wet winters or within small catchments). In such catchments, the predicted long-term mean flow may be biased if the rainfall error is significant.
- The long-term water balance, in all catchment types, will be affected by biased (as opposed to unbiased) rainfall error within the input data, though this will be less significant in areas of lower average annual rainfall (<700 mm).
- Low flow statistics and mean monthly flow predictions were largely insensitive to this type of error in all catchment types. This suggests that model predictions would still be useful to end-users who do not require accurate predictions of peak flows.

Model structural limitations – soil moisture deficits

- Structural limitations in model design that may result in over-estimation of summer soil moisture deficits (and thus under-estimation of streamflow response), will only significantly affect the accuracy of peak summer flow predictions within low BFI and wet catchments where higher rainfall and resulting streamflow would be expected.
- Model predictions of monthly mean flow and monthly flow variance are largely insensitive to this type of error in all catchments.

Parameterisation error – base flow drainage

- Parameter estimation errors that affect the rate of base flow drainage (i.e. errors in estimation of K_g), will alter model performance within all catchment types.
- Over-estimation of the K_g parameter in high BFI catchments leads to over-estimation of peak flows and under-estimation of the underlying baseflow signal in both summer and winter. This would lead to exaggerated estimations of flood risk potential within the spring–summer stream-recharge period, and under-estimate low flow levels within the late-summer to winter period.
- Under-estimation of the K_g parameter within low BFI catchments leads to an under-estimation of peak flows and an over estimation of low flows throughout the year.

Parameterisation error – recession curve

- Parameter estimation errors that affect the predicted rate of flow recession (i.e. errors in estimation of K_b) will affect model performance within all catchment types.
- Prediction of low flows within low BFI catchments were most sensitive to errors in parameterisation of the K_b parameter. This would make the use of such predictions in ecological studies difficult.
- Estimation of peak flows will be insensitive to changes in this parameter, making peak flow predictions in low BFI catchments more reliable.
- In higher BFI catchments predictions of both peak and low flows will be more sensitive.

The end use or utility of a time series will determine the what information is required, and its accuracy. Taking the simplest case, there is a lot more information in a time series of daily flows than a time series of monthly flows and a higher level of model performance is required to gain an accurate daily time-series simulation. This report has provided guidance on the sensitivity of different measures of utility to different aspects of predictive uncertainty in the modelling process. This guidance, in conjunction with the assessment of the sources of predictive uncertainty presented within Task 2.1, forms the basis of advice on whether the intended use of the time series information will be sensitive to potential sources of uncertainty and hence whether the simulation is fit for purpose.

The following deliverables have been produced in this task:

- Short report identifying likely sources of modelling error and the implication of such error to model implementation and use.

References

- Allen, R.G., Pereira, L.S., Raes, R. and Smith, M. 1998. *Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements*. Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, Rome, 300pp.
- Boorman, D.B., Hollis, J.M. and Lilly, A. 1995. *Hydrology of Soil Types: A Hydrologically-Based Classification of the Soils of the United Kingdom*. Institute of Hydrology Report No. 126. Wallingford, UK.
- Morris, D. and Heerdegen, R. 1988. Automatically derived catchment boundaries and channel networks and their hydrological applications. *Geomorphology*, 1, 131–141.
- Nash, J.E. and Sutcliffe, J.V. 1970. River flow forecasting through conceptual models, part 1: a discussion of principles. *Journal of Hydrology*, 10, 282–290.
- Smith, G.M., Fuller, R.M., Sanderson, J.M., Hill, R.A. and Thompson, A.G. 2001. Land Cover Map 2000: a parcel-based approach from satellite images. *Proceedings of the RSPS meeting in Uncertainty and Remote Sensing and GIS*, pp. 689–702.
- Young, A.R. 2006. Stream flow simulation within UK ungauged catchments using a daily rainfall–runoff model. *Journal of Hydrology*, 320(1–2), 155–172.
- Young, A.R., Holmes, M.G.R., Round, C. and Sekulin, A. 2000. *Project Summary: Enhanced Low Flow Estimation at the Ungauged Site and Modelling Historical Flow Sequences at the Ungauged Site*. CEH Wallingford and the Environment Agency.

List of abbreviations

AE	Actual evapotranspiration
BEQ95	Simulation bias error at Q95
BIAS	Simulation bias error
BFI	Baseflow index
CAMS	Catchment Abstraction Management Strategies
CEH	Centre for Ecology and Hydrology
CERF	Continuous Estimation of River Flows
CV	Coefficient of variance
DTM	Digital terrain model
EP	Effective precipitation
FAO	Food and Agriculture Organization of the United Nations
FC	Field capacity
HOST	Hydrology of Soil Types
HRU	Hydrological Response Unit
I_j	The interception depth within the day from vegetation class, j (mm)
I_p	Potential depth of rainfall that could be intercepted from the precipitation depth within the time step
K_b	Slow flow routing parameter: for a base flow linear reservoir (hours)
K_g	Drainage time constant: storage capacity distribution parameter (hours)
LF_OBJ	Low flows objective function
LCM2000	Land Cover Map 2000 (CEH)
MORECS	Met Office Rainfall and Evaporation Calculation System
MOSES-PDM	Met Office Surface Exchange Scheme with Probability Distributed Moisture
n	Number of cell increments in the pathway to outfall point
NSE	Nash-Sutcliffe model efficiency
p	Land-cover defined depletion factor
PE	Potential evapotranspiration
RAW	Readily available water
R^2	Square of the correlation coefficient (explained variance)
SEISMIC	Spatial Environmental Information System for Modelling the Impact of Chemicals

$Sh(i)$	Soil type specific velocity constant of cell i
SMD	Soil moisture deficit
TAW	Total available water
WFD	Water Framework Directive
WP	Wilting point
x_i	Actual distance between cell (i) and downstream cell (i+1)
Z_r	Root depth
$\bar{\beta}$	Average slope between cell i and cell (i+1)

We are The Environment Agency. It's our job to look after your environment and make it **a better place** – for you, and for future generations.

Your environment is the air you breathe, the water you drink and the ground you walk on. Working with business, Government and society as a whole, we are making your environment cleaner and healthier.

The Environment Agency. Out there, making your environment a better place.

Published by:

Environment Agency
Rio House
Waterside Drive, Aztec West
Almondsbury, Bristol BS32 4UD
Tel: 0870 8506506
Email: enquiries@environment-agency.gov.uk
www.environment-agency.gov.uk

© Environment Agency

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.