Qube Technical Note

Time Series Modelling

Qube Technical Note
Time Series Modelling

Document issue details

WHS8035

Version	Issue date	Issue status	Prepared By	Approved By
1	31/05/2019	Final	Prof. Andy Young (*Lead Scientist*)	Jude Jeans (*Director*)

For and on behalf of Wallingford HydroSolutions Ltd.

Contents

1 The Challenges of Estimating River Flow Time Series

Information on the magnitude and variability of flow regimes, at the river reach scale is a central component of most aspects of water resource and water quality management. For some activities, such as the setting of discharge consents and licencing of small abstractions, it is sufficient to encapsulate this information using a statistical description of the flow regime. Historically UK design procedures have focused on the estimation of the Flow Duration Curve (FDC). The initial focus was on the estimation of natural FDCs (NERC, 1980[1], Gustard et al. 1992[2]) and then, in 1994, these procedures were extended to the estimation of influenced FDCs. These early procedures were implemented within the Micro LOW FLOWS software (Young et al, 2000[3]). Use identifies gaps in methods and hence both methods for the estimation of natural and artificial influences have continued to evolve over the intervening years. Step changes in methods were delivered through the LowFlows 2000 software in the early 2000s (Holmes et al., 2002[4]) and latterly through the next generation of that software, the LowFlows Enterprise software (LFE).

Qube is the latest web-based software deployment of the LFE methods with significant improvements to the delivery of methods. However, the name Qube reflects that the design problem is not restricted to the estimation of FDCs; there are many applications for which a time series of river flows is required. These include the assessment of yield for water resource schemes, the in-stream flow requirements of aquatic flora and fauna and the assessment of the impacts of climate change at the catchment scale.

In estimating the time series of river flows at a site it is not necessary to exactly replicate all aspects of the true time series. Rather, the requirement is to simulate important facets of the regime including:

- an acceptable simulation of mean flow – conservation of mass;
- the correct statistical properties of the variation of river flow frequency distribution (best represented as the non-parametric FDC);
- how the stream flow reduces in the absence of rainfall - termed recession behaviour;
- the correct representation of seasonal patterns within the flow regime; and
- the correct stream flow response to precipitation and the dependencies of that response on antecedent catchment conditions.

With regard to the last point, it is not important to accurately simulate individual high flow events. The only restrictions on the modelling of high flows are that mass must 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 the high concentrations of suspended solids. As the cause and effect links between flow and habitat for aquatic species cannot be accurately quantified, predictive methods for assessing the ecological impacts of high flows are also not sensitive to the absolute magnitude of the flows.

---

[1] Natural Environment Research Council (NERC) 1980. Low Flow Studies report. Wallingford, UK.
[2] Gustard, A, Bullock, A. and Dixon, J.M. 1992. Institute of Hydrology Report 108. Low Flow Estimation in the United Kingdom
[3] Young, A.R., Gustard, A., Bullock, A., Sekulin, A.E., Croker, K.M. 2000. A river network based hydrological model for predicting natural and influenced flow statistics at ungauged sites, Science of the Total Environment. 251/252, 293-304.
[4] Holmes, M. G. R., Young, A. R., Goodwin, T. H. and Grew, R. 2005. A catchment-based water resource decision-support tool for the United Kingdom. Environmental Modelling & Software. Vol. 20, Issue 2, pp 197-202.

The overall design problem of estimating natural and influenced flows, expressed as flow statistics and time series and for both gauged and ungauged catchments is conceptualised as a cube, see Figure 1. The cube comprising 8 cube-lets, with each cube-let representing the river flow estimation requirement:

1. Gauged actual (influenced) time series.
2. Gauged actual (influenced) statistics.
3. Gauged naturalised (natural) time series.
4. Gauged naturalised (natural) statistics.
5. Ungauged natural time series.
6. Ungauged natural statistics.
7. Ungauged influenced (actual) time series.
8. Ungauged influenced (actual) statistics.

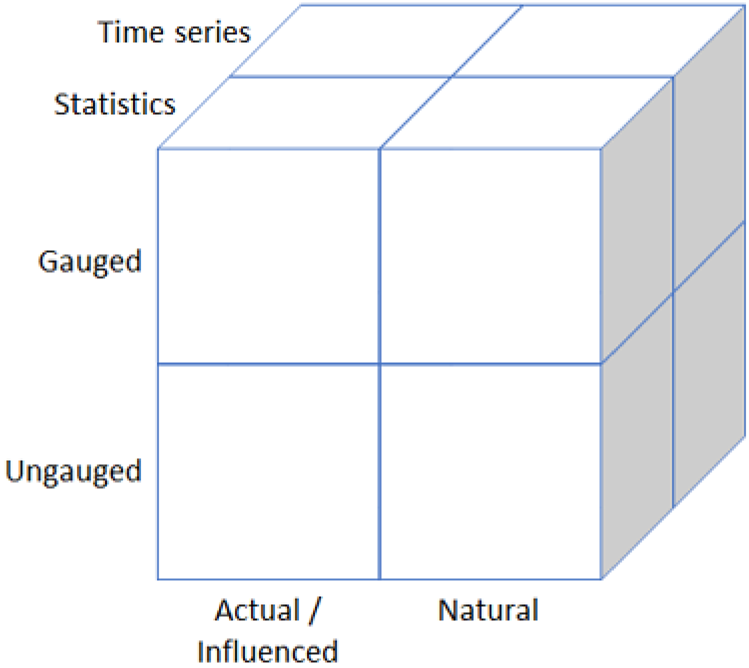

Figure 1 The Design Cube

Qube seeks to deliver all of this functionality in a consistent seamless way that:

- maximises the use of local gauged data within a river basin to constrain estimation uncertainty within ungauged sub-catchment; and
- provides a seamless transition between the estimation of natural and influenced FDCs and flow time series such that the flow time series preserves the statistical properties of the FDC.

Ignoring hydrometric error, the gauged case is simple, where the actual (or influenced) flow time series is that which is gauged and the actual FDC is the FDC derived from the gauged record. The industry best current practice for estimating the natural stream flow record from a gauged record is to naturalise the flow record through decomposition (Hall and Nott, 1994[5]; Young and Sekulin, 1996[6]). These methods were extended further through the Environment Agency's Practical Toolkit for Flow Naturalisation (Environment Agency[7]). Procedures for naturalising gauged river flows using the LFE influence data were implemented within the LFE software and have been re-engineered within Qube.

The ungauged catchment is more problematic. Research conducted alongside the development of the FDC estimation procedures within LowFlows 2000 developed generalised rainfall runoff models for application in the UK (Young, 2006[8]). The inherent problems with generalised lumped catchment models were further addressed by the development of the Continuous Estimation of River Flows (CERF) semi-distributed, generalised rainfall runoff model (Young et al., 2006[9]). CERF has been used successfully in many applications including climate change scenario modelling (Prudhomme, et al 2013[10]). A summary of the CERF model is provided in Appendix 1.

Notwithstanding this, the following problems with operational deployment of generalised rainfall runoff models persist:

- The software engineering challenges of managing and deriving the rainfall and potential evaporation input data for application within a catchment; and
- The fundamental limitation of a rainfall runoff model explaining the full variation within a dataset.

The former is a question of investment in software and hardware, the latter is more intractable. The problem is that a daily rainfall runoff model will tend to under estimate the highest flows and over estimate the lowest flows. This is a fundamental problem with any model, but it is exacerbated in a rainfall runoff model where the skill of the model is expected to both replicate the sequencing of flows and preserve the underlying statistical properties of the frequency distribution. Inevitably there is a trade off in both calibration and generalisation. In contrast, the FDC estimation procedures are focused on only estimating the statistical form of the flow distribution. It is unsurprising that in general the FDC estimation procedures provide a more accurate estimation of the FDC than the FDC derived from the outputs of a generalised model.

---

[5] Hall, J.K. and Nott, M.R. 1994. The naturalisation of flow records by decomposition. British Hydrological Society National Meeting on Flow Naturalisation Using Hydrological Models, London 17th March 1994.
[6] Young, A.R. and Sekulin, A.E. 1996. Naturalised River Flow Records of the Essex Region: Phase III Final Report. Institute of Hydrology unpublished Client Report.
[7]https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/290432/scho1005bjwz-e-e.pdf
[8] Young, A.R. 2006. Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. Journal of Hydrology, 320, 1-2, pp 155-172.
[9] Young, A.R., Keller V. & Griffiths J. 2006. Predicting low flows in ungauged basins; a hydrological response unit approach to continuous simulation. In Climate Variability and Chan- Hydrological impacts. International Association of Hydrological Sciences Publication 308. pp 134-138.
[10] Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., Kelvin, J., Mackay, J., Wang, L., Young, A.R and Watts, G. 2013. Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth Syst. Sci. Data, vol 5, pp 101-107.

2 Qube Time Series Estimation Method

2.1 The Basic Method

The requirement is for an operational method for estimating both natural and influenced flow time series for a site that are consistent with the corresponding best estimate of the FDC. The best estimate of the FDCs will be that estimated by Qube making maximum use of the available local data within the catchment.

The basic method is to sample the monthly FDCs from Qube using the best available Time Series of Exceedance flow Percentiles (TSEP) that can be estimated for the site. Depending on the availability of gauged flow data within the catchment these monthly flow percentile time series may be extracted from a gauged record that is judged to be a suitable donor. If a suitable gauging station is not available, then Qube will use the corresponding flow percentile time series extracted from a local CERF simulation.

The accuracy of the set of TSEPs for a gauging station is not sensitive to typical hydrometric errors which present as biases to estimation of low flows or high flows as only the rank of the flows is required. Thus, the pool of gauging stations suitable as TSEP sources is much larger than the corresponding pool of local data gauging stations. Similarly, the TSEPs are not sensitive to artificial influences other than influences that represent a significant fraction of the water balance and exhibit significant interannual or seasonal variability. Theoretically the naturalisation procedures within Qube could account for strongly seasonal influences, however these influences also tend to exhibit stronger interannual variability.

The gauging stations held on the National River Flows Archive were reviewed based on hydrometry (documented evidence of large changes in hydrometric quality) and the magnitude and likely seasonality of influences. The suitability of all NRFA gauging stations has been classified for this purpose as:

- Y - Natural or low net influence
- P - Significant net influence (but no impounding reservoirs)
- N - Unsuitable (stations with unstable hydrometric records (e.g. large changes to the measurement structure) and those that are influenced by large temporally varying influences).

All gauging stations classified as Y or P are considered as potential donor TSEP sites. Approximately 1,000 gauging stations have been selected across Great Britain. Where the gauged record is not complete, this is infilled by CERF simulated monthly exceedence percentiles data.

To augment the gauged TSEP sources, approximately 10,000 CERF simulations have been generated across Great Britain for potential donor TSEP sites. TSEP will be insensitive to overall bias (water balance error) and biases resulting in over or under-estimation at high or low flows. However, these CERF TSEPs will be sensitive to the incorrect simulation of the timing of flows. To take this into account in the methods, the rank correlation of CERF simulations was evaluated at all ROI gauging station. An average rank correlation was identified to reflect the loss of correlation associated with using a CERF simulation as a source of the TSEP rather than a gauging station.

Currently the Qube method applies to the full period of record from 1961 to the latest date at which the underlying time series were loaded.

2.2 Method Rationale

The method is best explored through consideration of a gauged record. If we consider a gauged natural time series of 30 years, and hence "n" daily flow values, where n =30 x 365 values (give or take a leap year). An annual FDC can be constructed by ranking (r) all the measured flows in order of decreasing size and then calculating the exceedance probability for each flow as r/(n+1). The FDC can then be graphed from this and a good representation can be obtained by abstracting 12*30 equally spaced Q(x), P(x) plotting positions.

We could then take the raw flow time series, assign an exceedence probability to each day within the time series based on the rank order approach above. This would create a time series of exceedance probabilities (TSEP).

We could then obtain a perfect time series by taking each exceedance probability in the time series and identifying the flows for the two FDC exceedance plotting positions that bracket the value and interpolating the flows between them accordingly. If the original full n plotting positions were retained in the FDC then this approach would return a perfect time series. Even with a reduced set of plotting positions (such as the 12 x 101 monthly FDC used within the Qube) the resultant time series is insignificantly different from the gauged.

Of course, one would never do this in practice at a gauging station. But in practice with an estimated TSEP at an ungauged catchment and using the corresponding estimate of the flow duration curves from Qube for the catchment time series of river flows for the site using this approach.

2.3 TSEPs Data Source and Period of Record

As discussed, the monthly TSEPs may be from measurements (gauging station) where the influence at the gauge is such that the actual TSEPs is unlikely to be different from the natural one. If a suitable gauging station is not available, then the monthly TSEPs are derived from a CERF simulation for a standard period from 1961 to the current last year of record in Qube.

The period of record data for individual gauging stations may not span this entire period and may have missing data within the record. To enable a full time series to be simulated missing data are infilled with CERF simulations. The extent of missing data at the beginning, end and within gauged period of record are recorded for use in estimating likely correlation, discussed in section 3. The CERF infilling procedure is a pragmatic one. Using the flow exceedance percentiles ensures the outcome is a CERF simulation which has the same frequency distribution as the gauged record.

3 Selecting a Donor TSEPs based upon Rank Correlation

The selection of a TSEP at a location where all or part of the TSEP is based upon a gauging station is obvious. Away from a TSEP gauging station the question is whether to use a TSEP from the location of a potentially distant gauging station or to use the TSEP from a "closer" CERF simulation recognising that the CERF simulation is an imperfect estimate of a natural TSEP for the location of the CERF simulation.

3.1 Measure of Suitability

The suitability of the time series for use will vary from catchment to catchment. The core method for selecting an appropriate TSEP site is based upon maximising the likelihood of replicating the timeseries information that would be observed if that location was gauged. Similarity is measured by Spearman Rank Correlation (SRC), which is calculated between two time series by considering the pearson correlation between two rank ordered time-series. A rank correlation of 1 is a perfect correlation of Ranks.

There are two rank correlations to be considered:

- The at site SRC – how well does the TSEP for a donor catchment represent the true TSEP for that catchment?
- What would be the SRC between a true TSEP for a donor catchment and the true TSEP for the target catchment that the TSEP is to be transferred to?

The combination of these two SRC values is the combined SRC value for a donor catchment.

3.1.1 At Site SRC

A rank correlation of 1 is a perfect correlation of ranks and thus a perfect TSEP estimate. So, for a TSEP at a site comprising a complete gauged record the at site SRC will be 1 as the time series is measured. If the TSEP is sourced from CERF, then the SRC will be less than one - no rainfall runoff model will yield a perfect simulation of gauged flows, and hence the rank order of flows. However, in the absence of gauged data to value the CERF simulation against the at site SRC will need to be estimated. This estimate is based upon the performance of CERF in similar catchments that are gauged. A value of 0.8 was selected based on an analysis of at-site SRCs obtained for CERF simulations across the Region of Influence pool of catchments used within Qube for flow duration curve estimation.

If the TSEP for a location is a composite of CERF and gauged records, then the CERF at site SRC will have been calculated directly between the gauged flows and the corresponding CERF flows for the days on which there are gauged flows. The TSEP for the infilled gauged record is then weighted between 1 for the proportion of gauged values and the CERF at site SRC for the proportion of CERF simulated values.

3.1.2 SIMINDEX

The second SRC value to be considered is the SRC between the true TSEP for the target site and the true TSEP for a donor site is estimated using the SIMINDEX regression model. SIMINDEX estimates the likely correlation between a distant donor source of time series data and a target site based on the distance between catchment outlets, BFIHOST and LFE RUNOFF. This model was developed using the SRC between pairs of ROI gauged catchments that met an overall proximity threshold of less than 50km. It does not differentiate between nested and unnested catchments as there were only 40 nested pairs within the dataset which was insufficient to construct a regression relationship.

The SIMINDEX relationship for England and Wales is given by:

$$SIMINDEX = 0.9857 - 0.0001689 * \Delta RUNOFF(1961-90) - 0.5276 * \Delta\_BFIHOST - 0.01047 * SQRT\ (DIST)$$

Where:

DIST = distance between catchment outlets;

RUNOFF(61-90) = the Qube estimate of natural runoff for this period; and

BFIHOST = The estimate of BFI derived from catchment soil classes.

SIMINDEX does not include the difference in catchment area as it wasn't significant statistically. Intuitive hydrological reasoning would suggest that the difference in area is important in determining the SRC between gauged catchments, as would nesting. The reason why it was not significant is that the influence of area is covariant with the influence of runoff and distance and a threshold of 50km separation was imposed on the selection of pairs for analysis. Nevertheless, to be prudent, area constraints have been incorporated into the selection of donor TSEP.

3.2 Selecting a TSEP Donor

The fundamental requirements of the time series estimation method are a correct estimate of the natural and influenced FDCs for a catchment and a representative TSEP estimated based on an appropriate donor source of flow percentile time series. This selection needs to differentiate between gauged and modelled percentile series.

TSEPs that have high at site SRC (SITE_SRC) values because they are dominated by gauged records will have a greater value in many instances than a TSEP primarily based on CERF. As the density of CERF based TSEPs is much greater than ones dominated by gauged records, there is only the requirement to consider the first order nested CERF TSEPs. These are the first one downstream and the upstream ones than are non -nested with each other (i.e. the first ones that are encountered climbing up all branches of the upstream drainage network).

For selecting TSEPs classified as predominately gauged a looser set of criteria are used. Candidate TSEP donors are selected within 50km and area factor 4, based on a hierarchy classified by TSEP type and whether they are nested or adjacent to the target. More stringent rules are applied to adjacent sites or gauging stations with a significant influence and small catchments (less than 20 km^2).

This selection process yields a list of n TSEP candidates. For each candidate donor catchment the product of SIMINDEX and SITE_SRC is calculated and Qube selects the TSEP with the highest product value. The proof for combining SIMINDEX and SITE_SRC in this way is given in Appendix 2.

4 Sampling from Flow Duration Curves to create Time Series

Having selected the donor TSEP, the next step is to select the appropriate monthly FDC set to sample from. This will depend on whether the donor TSEP is:

- CASE 1: predominantly CERF based, and thus assumed to be natural or whether it is adjacent gauged, in which case the assumption has to be made that it is natural (obviously this may not be the case in practice).
- CASE 2: downstream or upstream gauged, in which case it is assumed to be influenced and the influences that may affect the TSEP are nested to both the gauged and target catchments this may not be the case in practice also the loss of information if this is not correct is likely to be small.

The FDC estimation procedures for each case are as follows.

CASE 1: Assumed Natural TSEP

In this case the TSEP are used to sample from the corresponding natural FDCs using simple log-linear interpolation between adjacent exceedance probability plotting positions.

The influenced time series is then estimated by interpolating between the two influenced flows corresponding to the adjacent natural FDC p(x) value to the p(x) value within the selected source of time series information.

CASE 2: Assumed Influenced TSEP

In this instance the natural time series is estimate by interpolating between the natural flows corresponding to the two adjacent influenced FDC p(x) for each MTSEP series and corresponding influenced FDC.

The influenced time series by interpolating between the two adjacent influenced flows corresponding to the adjacent influenced FDC p(x) values for the month in question.

Appendix 1 An Overview of the CERF model

The CERF model uses a flexible hydrological response unit based model structure in which catchment descriptors of vegetation and soil type are used to define more complex, specific model structures within each catchment than can be identified from stream flow data from a single catchment. To enable subsequent parameter identification from streamflow data, the model parameters are simultaneously optimised against streamflow data across all catchments to combine model calibration and regionalisation in a single step process.

The catchment descriptors used within the model structure are:

- a hydrologically referenced 50m resolution, Digital Terrain Model (DTM) (Morris & Heerdegen, 1988[11]);
- the Hydrology of Soil Types (HOST) 29 class hydrological response classification of soils across the United Kingdom (Boorman et al., 1995[12]); 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 50m resolution (Smith et al., 2001[13]).

The model structure of CERF (based around Hydrological Response Units) is illustrated within Figure 2. The model structure is based around two sub model components presented diagrammatically;

- 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[14]). There is also an interception model (not depicted in Figure 2) which was 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 module describes vegetation as a function of maximum root depth, Zr, 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 Zr. Plants freely transpire until Soil Moisture Deficits (SMD) exceed the threshold defined by p.Zr, 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.

---

[11] Morris, D. and Heerdegen, R. (1988). Automatically derived catchment boundaries and channel networks and their hydrological applications. Geomorphology, 1, 131-

[12] 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.

[13] 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, p689-702.

[14] 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, 300 p.

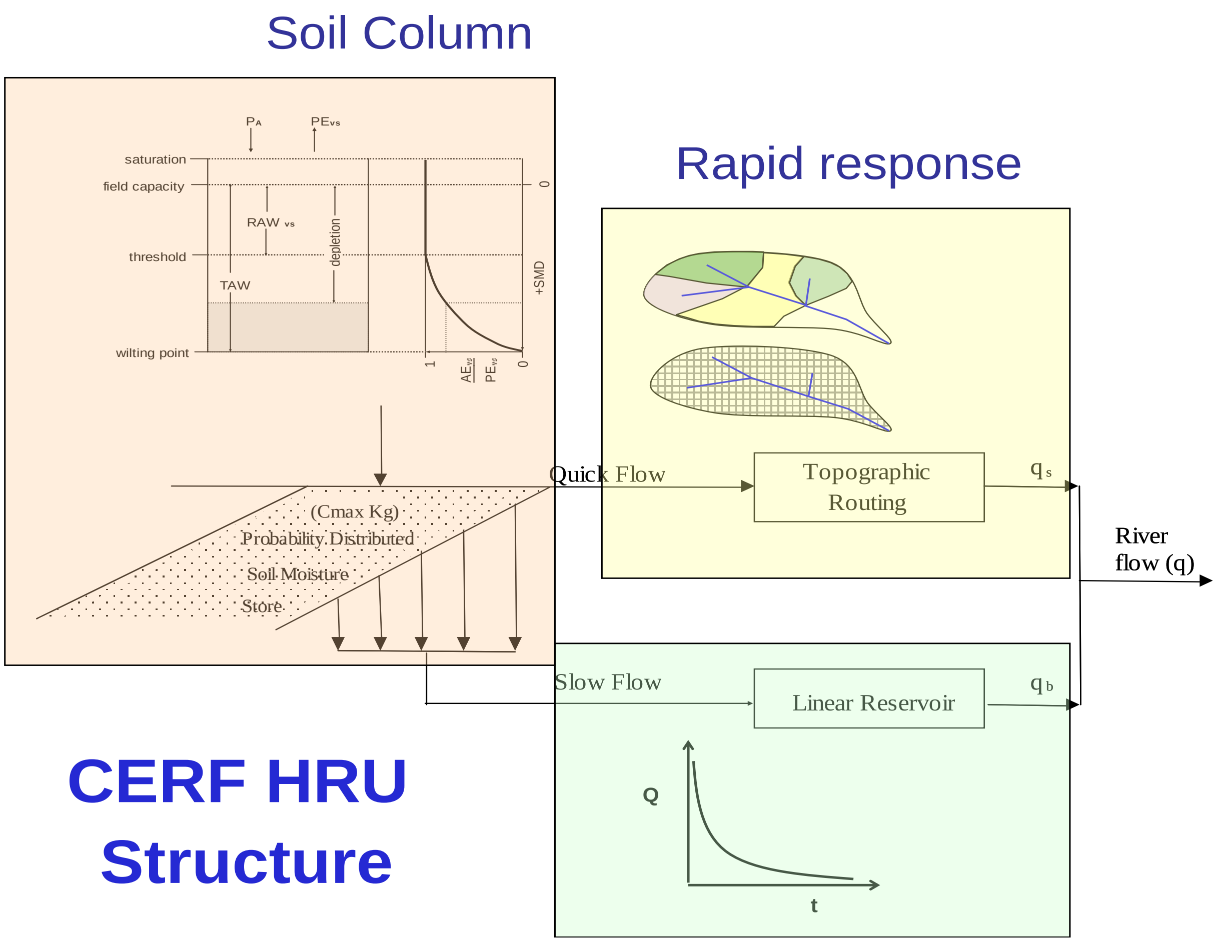

Figure 2 The CERF model structure

HRUs are 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. In practice the number of actual combinations is significantly less as some land-cover/soil class combinations do not occur. 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, Zr 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 data set. This process leaves Zr and p, for each vegetation class as the free parameters for the HRUs within the loss module which equates to 10 parameters in total.

The output from the loss module within a catchment is an EP time series for each 1km 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 influences on the routing of water through the land surface are soils, hydrogeology and topography. 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. A probability distributed storage model representing free water in the soil column for the routing HRU. This storage model is assumed to be uniformly distributed with a maximum storage depth of 75mm (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, Kg, is a free parameter. The runoff from the store is routed through a topographically defined routing path, whilst the drainage is routed through a linear reservoir, with a time constant Kb, representing the baseflow for the routing class.

The quick flow routing within the routing HRUs is 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 based upon the flow path defined from the DTM and the cell level topographic gradients along the path, □. Total travel time to the catchment outlet, T, is calculated for each cell as

$$T = \sum_{i=1}^{N} \frac{X_i}{\beta_i v} , \qquad (1)$$

where X_i is the distance between the centroids of adjacent cells within the flow path and v is a routing HRU dependent velocity 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 of time constant, Kq, to represent the transient storage along the flow path lengths. For each routing HRU the model has four free parameters; Kg, V, Kb and K1 and with 11 routing class this gives a total of 44 routing parameters within the model which combines with the 10 free loss module HRU parameters yields a total of 54 model parameters for calibration.

Within the original research project, the parameters were simultaneously optimised across all catchments within the calibration sets using data from the 15 year period from 1987-2001. This process combines both model calibration and regionalisation into a one step process as the model parameters are a function of the HRUs and hence the descriptors of the catchments. The loss module HRUs were initially calibrated to make the process of simultaneous calibration computationally tractable. The mean and variance of the distribution of the bias between simulated mean EP and the corresponding observed mean runoff, expressed as a percentage of the observed mean runoff (BIAS) over the calibration period was minimised as the objective function. The routing module parameters were subsequently optimised based on the trade off between maximising the Nash Sutcliffe Efficiency Criterion (N.S.E.) and minimising the sum of squared deviations between observed and simulated stream flow over the lowest third of the flow distribution (LF_OBJ) and the bias error at the Q95 flow (BEQ95) across all catchments within the calibration dataset. N.S.E. was used as a general measure of fit whilst the latter functions are a measure of fit at low flows. The earlier periods of flow data within the calibration catchments and the evaluation catchments were used to ensure the model was not over fitted within the calibration catchments.

Appendix 2 Combining Correlations

The proof for combining correlations is as follows. Consider a target site A and a donor site B. Consider that the true TSEPs for A and B have a SRC of x. Consider we have an estimate of the true time series for B, call it C, that has a correlation of y with the true time series at B. Now consider correlations x y (so correlation between AB is x and correlation between BC is y) then the minimum correlation z between and A and C is:

z_{min} = xy -sqrt(1-x^2)sqrt(1-y^2)

Where the terms (1-x^2) and (1-y^2) are those parts of the correlations between AB and BC that are not part of the correlation of AC (the orthogonal components).

If we make the assumption that the correlation of C with B is quite comparable with the correlation one would have with a true time series at A if method C was applied at A. In this case then there are no orthogonal correlation components between AB and BC and z_{max}=xy.

This is a reasonable approximation to make and introducing empirical weights to optimise the orthogonal components is likely to result in an overfitted model.

In the table below the z_{max} values are intuitively reasonable.

x	y	z_{min}	z_{max}
0.9	0.8	0.46	0.72
0.9	0.7	0.32	0.63
0.9	0.6	0.19	0.54
0.9	0.5	0.07	0.45
0.8	0.8	0.28	0.64
0.8	0.7	0.13	0.56
0.8	0.6	0.00	0.48
0.8	0.5	-0.12	0.40
0.7	0.8	0.13	0.56
0.7	0.7	-0.02	0.49
0.7	0.6	-0.15	0.42
0.7	0.5	-0.27	0.35
0.6	0.8	0.00	0.48
0.6	0.7	-0.15	0.42
0.6	0.6	-0.28	0.36
0.6	0.5	-0.39	0.30
0.5	0.8	-0.12	0.40
0.5	0.7	-0.27	0.35
0.5	0.6	-0.39	0.30
0.5	0.5	-0.50	0.25