Reversing Hydrology: quantifying the temporal aggregation effect of catchment rainfall estimation using sub-hourly data

Quantifying temporal aggregation effects of catchment rainfall estimation

Authors

Kretzschmar, A[1]., Tych, W[1*]., Chappell, N.A[1]., Beven, K.J.[1, 2]

[1] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

[2] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Corresponding author:

Dr. W. Tych,

Lancaster Environment Centre,

Lancaster University,

Lancaster

LA1 4YQ,

UK

w.tych@lancaster.ac.uk

+44 (0)1524 593973

Abstract

Inferred rainfall sequences generated by a novel method of inverting a continuous time transfer function show a smoothed profile when compared to the observed rainfall however streamflow generated using the inferred rainfall is almost identical to that generated using the observed rainfall ($R_t^2 = 95\%$). This paper compares the

inferred effective and observed effective rainfall in both time and frequency domains in order to confirm that, by using the dominant catchment dynamics in the inversion process, the main characteristics of catchment rainfall are being captured by the inferred effective rainfall estimates. Estimates of the resolution of the inferred effective rainfall are found in the time domain by comparison with aggregated sequences of observed effective rainfall, and in the frequency domain by comparing the amplitude spectra of observed and inferred effective rainfall. The resolution of the rainfall estimates is affected by the slow time constant of the catchment and the rainfall regime, but also by the goodness-of-fit of the model, which incorporates the amount of other disturbances in the data.

Keywords

Continuous time; data based mechanistic modelling; time resolution; spectral analysis; reverse hydrology; transfer function;

Introduction

Rainfall is the key driver of catchment processes and is usually the main input to rainfall-streamflow models. If the rainfall and/or streamflow data used to identify or calibrate a model are wrong or disinformative, the model will be wrong and cannot be used to predict the future with any certainty. Bloeschl *et al*. (2013) state that if the dominant pathways, storage and time-scales of a catchment are well defined then a model should potentially reproduce the catchment dynamics under a range of conditions. It is often the case that hydrological variables, such as rainfall and streamflow, are measured at hourly or sub-hourly intervals then aggregated up to a coarser resolution before being used as input to rainfall-streamflow models resulting in the loss of information about the finer detail of the catchment processes (Littlewood and Croke, 2008; Littlewood *et al*., 2010; Littlewood and Croke, 2013).

Kretzschmar *et al*. (2014) have proposed a method for inferring catchment rainfall using sub-hourly streamflow data. The resulting rainfall record is smoothed to a coarser resolution than the original data but should still retain the most pertinent information.

This paper investigates the implications of the reduced resolution and the potential loss of information introduced by the regularisation process in both the time and frequency domains. Both temporal and spatial aggregation are incorporated in the transfer function model however only the temporal aspect is considered here. The effect of spatial rainfall distribution using sub-catchments will be the subject of a future publication.

The method developed and tested by Kretzschmar *et al*. (2014) – termed the RegDer method - inverts a continuous-time transfer function (CT-TF) model using a regularised derivative technique to infer catchment rainfall from streamflow with the aim of improving estimates of catchment rainfall arguing that a model that is well-fitting and invertible is likely to be robust in terms of replicating the catchment system.

The classical approach to inverse (as opposed to reverse) modelling involves the estimation of non-linearity (rainfall or baseflow separation) and the unit hydrograph (UH), which is an approximation to the impulse response of the catchment. Boorman (1989) and Chapman (1996) use sets of event hydrographs to estimate the catchment UH. Boorman (1989) superimposed event data before applying a separation technique and concluded that the data required may be more coarsely sampled than might be expected because one rain-gauge is unlikely to be representative of the whole catchment. Chapman (1996) used an iterative procedure to infer rainfall patterns for individual events before applying baseflow separation.

The resultant UHs had higher peaks and shorter rise times and durations than those obtained by conventional methods. He viewed the effective rainfall as the output from a non-linear store. Duband *et al.* (1993) and Olivera and Maidment (1999) used deconvolution to identify mean catchment effective rainfall which was redistributed using relative runoff coefficients whilst Young and Beven (1994) based a method for inferring effective rainfall patterns on the identification of a linear transfer function. A gain parameter, varying with time accounted for the non-linearity in the relationship between rainfall and streamflow.

In recent years, a range of different approaches has been used to explore reverse modelling in hydrology, that is, estimating effective rainfall from streamflow. Notable publications include Croke (2006), Kirchner (2009), Andrews and Croke (2010), Young and Sumislawska (2012), Brocca *et al*. (2013, 2014) and Kretzschmar *et al.* (2014). Kirchner's method links rainfall, evapo-transpiration and streamflow through a sensitivity function making assumptions which allow rainfall to be inferred from the catchment streamflow. The method has been applied by Teuling *et al.* (2010) and Krier *et al*. (2012) to catchments in Switzerland and Luxembourg and has been found to work for catchments with simple storage-streamflow relationships and limited hysteresis. Brocca et al. (2013) employed a similar method based on the water balance equation but inferred the rainfall series from soil moisture. In a further study, Brocca et al. (2014) used satellite derived soil moisture to infer global rainfall estimates. Croke (2006) derived an event-based unit hydrograph from streamflow alone but his approach was limited to ephemeral quick-flow-dominant catchments whilst Andrews *et al*. (2010) and Young and Sumislawska (2012) use a discrete model formulation inverted directly or via a feedback model (which could be adapted to CT formulation). The approach proposed by Kretzschmar *et al*. (2014) combined a continuous time transfer function (CT-TF) model with regularized derivative estimates to infer the catchment rainfall from sub-hourly streamflow data.

Littlewood (2007) applied the IHACRES model (e.g. Jakeman *et al*.,1990) to the River Wye gauged at Cefn Brwyn showing that the values for the model parameters for that catchment changed substantially as the data time step used for model calibration decreased. Littlewood and Croke (2008) extended this work to include a second catchment and found that as the time-step decreased the parameter values approached an asymptotic level (on a semi-log plot) concluding that, at small enough time-steps, parameters become independent of the sampling interval. They suggested further investigation using data-based mechanistic modelling (DBM) methods as described by Young and Romanowicz (2004) and Young and Garnier (2006) for estimating CT models from discrete input data. Such models generate parameter values independent of the input sampling rate – as long as the sampling rate is sufficiently high in comparison to the dominant dynamics of the system. Advantages of using the CT formulation include allowing a much larger range of system dynamics to be modelled e.g. 'stiff' systems that have a wide range of time-constants (TC), typical of many hydrological systems. The outputs from such a model can be sampled at any time-step, including non-integer, and the parameters have a direct physical interpretation (Young, 2010).

Krajewski *et al*. (1991) compared the results from a semi-distributed model and a lumped model and concluded that catchment response is more sensitive to rainfall resolution in time than space whilst a study by Holman-Dodds *et al*. (1999) demonstrated that models calibrated using a smoothed rainfall signal (due to coarse sampling) may result in under-estimation of streamflow. Further calibration, required to compensate, leads to the loss of physical meaning of parameters. They also concluded that parameters estimated at one sampling interval were not transferable to other intervals; a conclusion echoed by Littlewood (2007) and Littlewood and Croke (2008).

Studies by Clark and Kavetski (2010) showed that in some cases, numerical errors due to the time-step are larger than model structural errors and can even balance them out to produce good results. The follow-up study by Kavetski and Clark (2010) looked at its impact on sensitivity analysis, parameter optimisation and Monte Carlo uncertainty analysis. They concluded that use of an inappropriate time step can lead to erroneous and inconsistent estimates of model parameters and obscure the identification of hydrological processes and catchment behaviour. Littlewood and Croke (2013) found that a discrete model using daily data over-estimated time-constants for the River Wye gauged at Cefn Brwyn when compared to those estimated from hourly data confirming that parameter values were dependent on the time-step. They discussed the loss of information due to the effect of time-step on time constants and suggested that plots of parameter values against time step could be used as a model assessment tool. In a previous study, Littlewood and Croke (2008), compared the sensitivity of parameters for two catchments with respect of time-step and discussed the role of time-step dependency on the reduction of uncertainty. They also suggested continuous time transfer function modelling using sub-hourly data to derive sampling rate independent parameter values. Littlewood *et al*. (2010) introduced the concept of the Nyquist-Shannon (N-S) sampling theorem, which defines the upper bound on the size of sampling interval required to identify the CT signal without aliasing, and consequentially its effect on the frequency of sampling required to specify a rainfall-streamflow model. Given a frequent enough sampling rate, the CT model is time independent and can be interpreted at any interval.

Further understanding may be gained by transforming rainfall and streamflow series from the time domain to the frequency domain and using spectral analysis. Several potential uses of spectral analysis in hydrology have been explored including modelling ungauged catchments, modelling karst systems and seasonal adjustment of hydrological data series. A maximum likelihood method for model calibration based on the spectral density function (SDF) has been suggested by Montanari and Toth (2007). The SDF can be inferred from sparse historic records in the absence of other suitable data making it a potentially useful tool for modelling ungauged catchments. They also suggest that spectral analysis may provide a means of choosing between different apparently behavioural models. Cuchi *et al.* (2014) used 'black box' modelling and frequency analysis to study the behaviour of a karst system (located at Fuenmajor, Huesca, Spain). They concluded that method works well for a linear system and that Fuenmajor has a linear hydrological response to rainfall at all except high frequencies. They suggest that the non-linearity issues might be addressed using appropriate techniques such as wavelets or neural networks. Szolgayova *et al.* (2014) utilised wavelets to deseasonalise a hydrological time-series and suggested that the technique had potential for modelling series showing long term dependency (interpreted as containing low frequency components).

The method introduced by Kretzschmar *et al.* (2014) showed that given that the rainfall-streamflow model captures the dynamics of the catchment system, the high frequency detail of the rainfall distribution is not necessary for the prediction of streamflow due to the damping (or low-pass filter) effect of the catchment response. The regularisation process introduced is numerically stable at the cost of a loss of some temporal resolution in the inferred rainfall time series. The regularisation level is controlled through the Noise Variance Ratio (NVR), optimised as part of the process and is only applied when necessary, i.e. when the analytically inverted catchment transfer function model is improper (has a numerator order higher than the denominator order).

Application catchments

RegDer has been tested on two headwater catchments with widely differing rainfall and response characteristics – Baru in humid, tropical Borneo and Blind Beck, in humid temperate UK. The 0.44 km^2 Baru catchment is situated in the headwaters of the Segama river on the northern tip of Borneo, East Malaysia. The climate is equatorial showing no marked seasonality but tending to fall in short (<15 min) convective events showing high spatial variability and intensities much higher than those of temperate UK (Bidin and Chappell, 2003, 2006). Haplic alisols, typically 1.5 m in depth and with a high infiltration capacity (Chappell *et al.*, 1998) are underlain by relatively impermeable mudstone bedrock resulting in the dominance of comparatively shallow sub-surface pathways in this basin (Chappell *et al.*, 2006). As a result of the high rainfall intensity and shallow water pathways the stream response is very flashy. In contrast, the Blind Beck catchment has an area of 8.8 km^2 and its response shows evidence of deep hydrological pathways due to the presence of deep limestone and sandstone aquifers resulting in a damped hydrograph response (Mayes *et al*., 2006; Ockenden and Chappell, 2011; Ockenden *et al*.,2014). Winter rainfall in this basin is derived from frontal systems with typically lower intensities than the convective systems in Borneo (Reynard and Stewart, 1993).

Model formulation and physical interpretation

This study investigated the limits of inferred catchment effective rainfall estimation from streamflow. Continuous time transfer function models identified from the observed data using Data Based Mechanistic (DBM) modelling approaches (Young and Beven, 1994; Young and Garnier, 2006), are inverted using the RegDer method (Kretzschmar *et al.*, 2014) and used to transform catchment streamflow into estimates of catchment inferred rainfall.

DBM modelling makes no prior assumptions about the model structure (though it often uses structures based on transfer functions), which is suggested by the observed data, and must be capable of physical interpretation. As transfer functions are linear operators, a transform structured as a bilinear power-law (Eq. (1)), also identified from the observed data, was applied to linearise the data before model fitting (Young and Beven, 1994; Beven, 2012, .p91):

$$\boldsymbol{P_e} = \boldsymbol{P}\,\boldsymbol{Q}^{\propto} \qquad (1)$$

where P is the observed rainfall, Q the observed streamflow and α is a parameter, estimated from the data. P_e is the effective observed rainfall (ER) and Q is used as a surrogate for catchment wetness. Both catchments used in this study proved to be predominantly linear in their behaviour so transformation Eq. (1) was not used. In the initial study, a wide range of possible models was identified using algorithms from the Captain Toolbox for Matlab (Taylor *et al.*, 2007). The models selected were a good fit to the data and were suitable for inversion. The Nash-Sutcliffe Efficiency (NSE or R_t^2) is commonly used to compare the performance of hydrological models. Often several models can be identified that fit the data well (the equifinality concept of Beven, 2006). From these, models with few parameters to be estimated that inverted well were selected. In this study a second order linear model was found to fit both catchments. The output from the RegDer process is an inferred effective rainfall series to which the reverse of the power law is then applied, if necessary, to construct an inferred catchment rainfall sequence. The process is illustrated in Fig. 1.

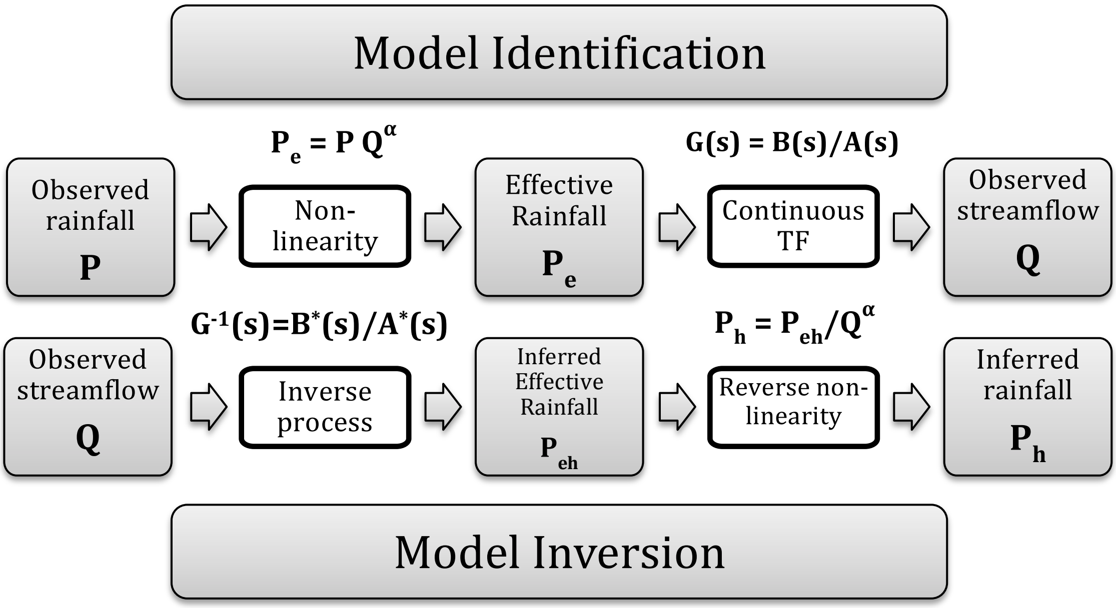


Figure 1 - model identification and inversion workflow where *P* is the observed catchment rainfall, P_e is the effective rainfall, *Q* is the observed streamflow, P_{eh} is the inferred effective rainfall and P_h the inferred catchment rainfall. Non-linearity is represented by the bilinear power law (Beven, 2012, p91). The continuous time transfer function is given by G(s) where A(s) and B(s) are the denominator and numerator polynomials and the inversion process is represented by $G^{-1}(s)$ where $A^{*}(s)$ and $B^{*}(s)$ are the denominator and numerator polynomials of the inverted transfer function.

The transfer function model inversion process has been described in Kretzschmar *et al.* (2014). It involves transition from the transfer function catchment model:

$$Q = G(s)R = \frac{\beta_0 s^m + \beta_1 s^{m-1} + \dots + \beta_m}{s^n + \alpha_1 s^{n-1} + \dots + \alpha_n} e^{-s\tau} P_e \qquad (2)$$

to the direct inverse (in general non-realisable):

$$\hat{\boldsymbol{R}} = \frac{\boldsymbol{b_0 s^n + b_1 s^{n-1} + \dots + b_n}}{\boldsymbol{s^m + a_1 s^{m-1} + \dots + a_m}} \boldsymbol{e^{s\tau} Q} \qquad (3)$$

which is then implemented using regularised streamflow derivatives in the form of:

$$\hat{\boldsymbol{R}}\boldsymbol{e^{-s\tau}} = \frac{\boldsymbol{b_0 \{\widehat{s^n Q}\}^{*} + b_1 \{\widehat{s^{n-1} Q}\}^{*} + \dots + b_n Q}}{\boldsymbol{s^m + a_1 s^{m-1} + \dots + a_m}} \qquad (4)$$

where $\{\widehat{s^n Q}\}^* = \mathcal{L}\left\{\widehat{\frac{d^n}{dt^n}Q}\right\}$ is the Laplace transform of the optimised regularised estimate of the n^{th} time derivative of Q: $\frac{d^n}{dt^n}Q$. The regularised derivative estimates replace the higher order derivatives in Eq. (3), which otherwise make Eq. (3) unrealisable (improper) – this is the core of the method in Kretzschmar *et al.* (2014). In the implementation, n^{th} derivatives in Eq. (4) are not estimated, but advantage is taken of the filtering with the denominator polynomial, and only $(n\text{-}m)^{th}$ derivative estimates are required in combination with a proper transfer function.

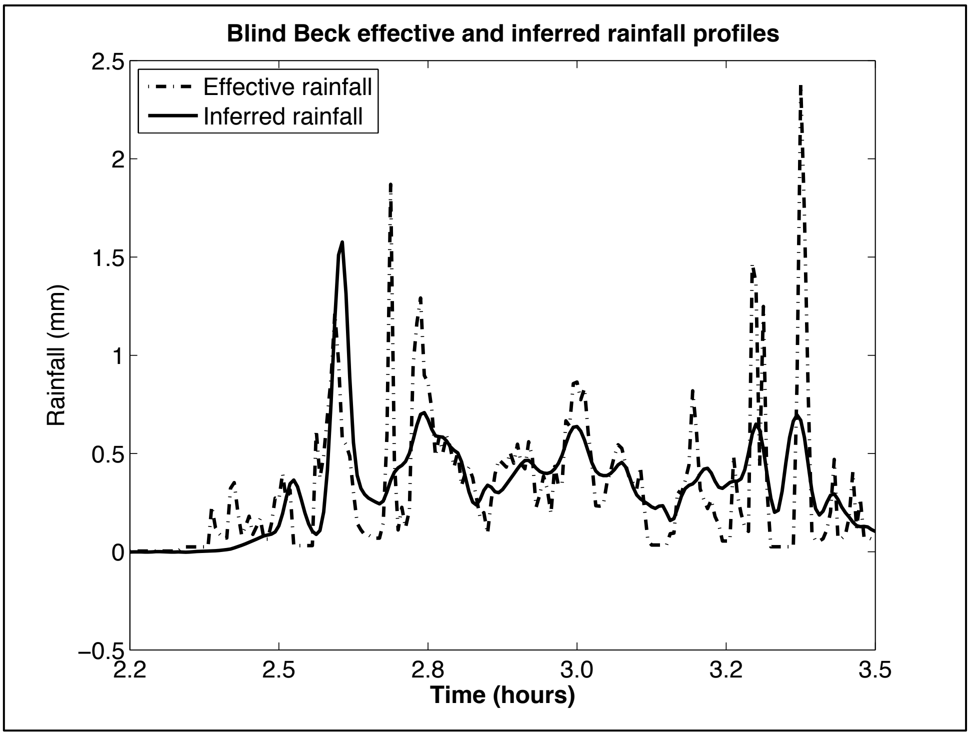
The inferred effective rainfall (IR) sequences generated by RegDer generally have a much smoother profile (illustrated in Fig. 2) than the observed rainfall inputs, however streamflow sequences generated with the IR used as the model input are almost indistinguishable from the sequence modelled using observed rainfall (R_t^2 = 95%). This indicates that the catchment dynamics, as captured by the transfer function model, renders the differences between observed and inferred rainfall immaterial. The reason for this becomes clear when looking at the frequency domain analysis of the inversion process shown in this paper.

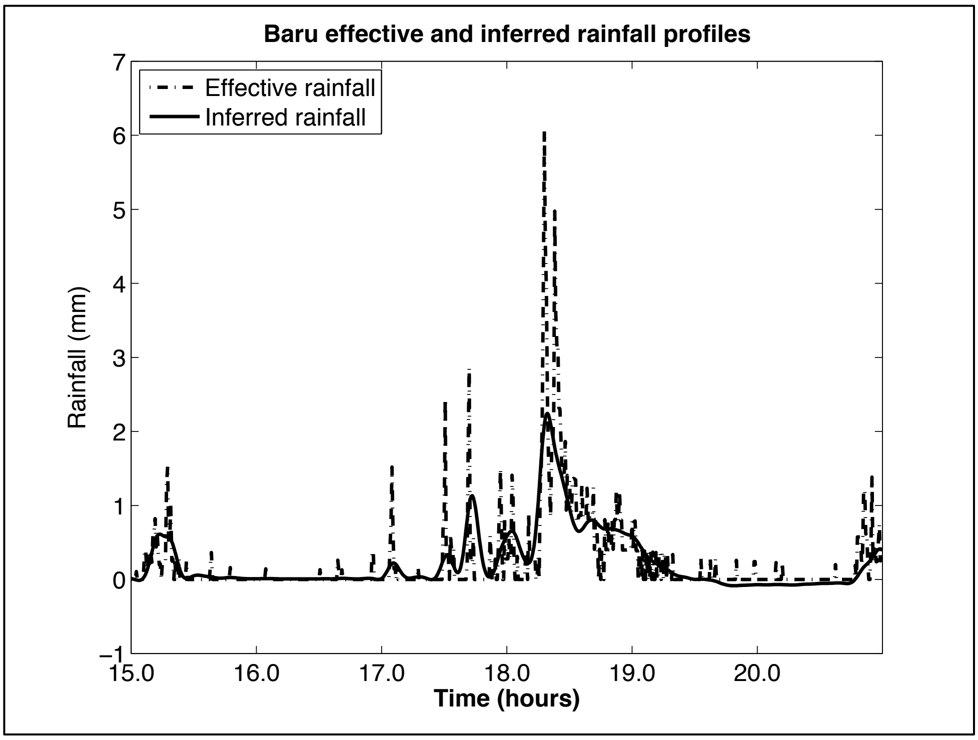
In order to investigate this, the IR is compared to aggregated effective observed rainfall sequences with increasing levels of aggregation until a good match is found (high value of R_t^2 or R). Two methods of aggregation have been used: 1) averaging over a range of time-series, 2) moving average over varying time scales. Two measures are used to assess the correspondence between the IR and the aggregated effective rain: 1) R_t^2, the coefficient of determination, and 2) R, the instantaneous Pearson correlation coefficient. They are given by:

$$R_t^2 = 1 - \frac{\sum(ER-IR)^2}{\sum(ER-\overline{ER})^2} \quad (5a)$$

$$R = \frac{\sum(ER-\overline{ER})(IR-\overline{IR})}{\sqrt{\sum(ER-\overline{ER})^2}\sqrt{\sum(IR-\overline{IR})^2}} \tag{5b}$$

where ER indicates a value from the aggregated effective rainfall sequence with mean $\overline{ER}$ and IR is the corresponding value from the inferred effective rainfall sequence with mean $\overline{IR}$. Both R_t^2 and R values tend towards a maximum value as aggregation increases. The aggregation level at which the maximum is reached is identified and taken as an estimate of the resolution of the inferred effective series. This value is then compared to the system fast time constant (TC_q) and the Nyquist-Shannon (N-S) sampling limit.

a) 



b)

Figure 2 – observed effective and inferred rainfall profiles generated using the RegDer inversion method for a) Blind Beck and b) Baru

Continuous model formulation

One of the advantages of using a CT model formulation is that the parameters have a direct physical interpretation independent of the model's sampling rate (Young, 2010). The continuous time model formulation for a 2nd-order model is given by:

$$y(t) = \frac{\beta_0 s + \beta_1}{s^2 + \alpha_1 s + \propto_2}\, u(t-\delta) \qquad (6)$$

where y is the measured streamflow at time *t,* δ is the transport delay and *u* is the effective rainfall at time *t* - δ. If the denominator can be factorized and has real roots, Eq. (6) can be rewritten as:

$$y(t) = \frac{\beta_0 s + \beta_1}{(s + \frac{1}{TC_q})(s + \frac{1}{TC_s})}\, u(t-\delta) \qquad (7)$$

where TC_q and TC_s are the system time constants and are often significantly different – a 'stiff' system. Decomposing the model into a parallel form gives:

$$y(t) = \left(\frac{g_q}{1+TC_q s} + \frac{g_s}{1+TC_s s}\right)u(t-\delta) \quad (8)$$

where g_q and TC_q are the steady state gain and time constant of the fast response component and g_s and TC_s are the steady state gain and time constant of the slow response component. The steady state gain of the system as a whole is given by:

$$g = g_q + g_s \quad (9)$$

so the fraction of the total streamflow along each pathway can be calculated from:

$$P_q = \frac{g_q}{g_q + g_s};\ P_s = \frac{g_s}{g_q + g_s} \quad (10)$$

The fraction of streamflow attributed to the slow response component is sometimes termed the Slow Flow Index (SFI) (Littlewood *et al.*, 2010). The example shown here uses a second order model but the general principle can be extended to higher order models. Details of the inversion and regularisation processes can be found in Kretzschmar *et al.* (2014).

Sampling frequency

When using CT modelling, the Nyquist-Shannon frequency gives the upper limit on the size of the sampling interval, Δt, that will enable the system dynamics to be represented without distortion (aliasing - Bloomfield, 1976, p21). Aliasing occurs when a system is measured at an insufficient sampling rate to adequately define the signal from the data.

The Nyquist-Shannon theorem states that the longest sampling step for a signal with bandwidth Ω (maximum frequency, where $\Omega = 2\pi f$ in cycles per time unit) to be represented is:

$$\Delta t \leq \frac{1}{2\Omega} \quad (11)$$

in order to completely define the system in absence of observation disturbance (Young, 2010). If the sampling interval is small enough to uniquely define the system, the estimated CT model should be independent of the rate of sampling. Conversely, if the frequency of the inferred output is less than the N-S limit, then the system dynamics should be adequately captured. Other estimates of the sufficient sampling interval, designed to avoid proximity to the Nyquist limit, have been made by Ljung (1999) and Young (2010). In terms of system TCs, these limits are given by:

$$Nyquist = \pi TC_q \ time\ units \tag{12a}$$

$$Ljung = \frac{\pi TC_q}{5} \ time\ units \tag{12b}$$

$$Young = \frac{TC_q}{6} \ time\ units \tag{12c}$$

Temporal aggregation of effective rainfall

Two methods for aggregating ER were used to estimate the time resolution of the IR. Rainfall is the total volume accumulated over the sampling interval so the ER was aggregated over progressively longer sampling periods of 2 to 24 times the base sampling period and averaged to form a new smoothed sequence that could be compared with the IR. For comparison, aggregation was also performed via a moving average process utilising the convolution method available in Matlab. Both methods may be affected by the aggregation starting point and edge effects. The aggregated ER sequences were compared to the IR using the coefficient of determination (R_t^2) and the correlation (R). R_t^2 and R tend towards a maximum value as aggregation increases. The aggregation time-step at which this value is established is used to estimate the resolution of the *IR*.

Spectral Analysis

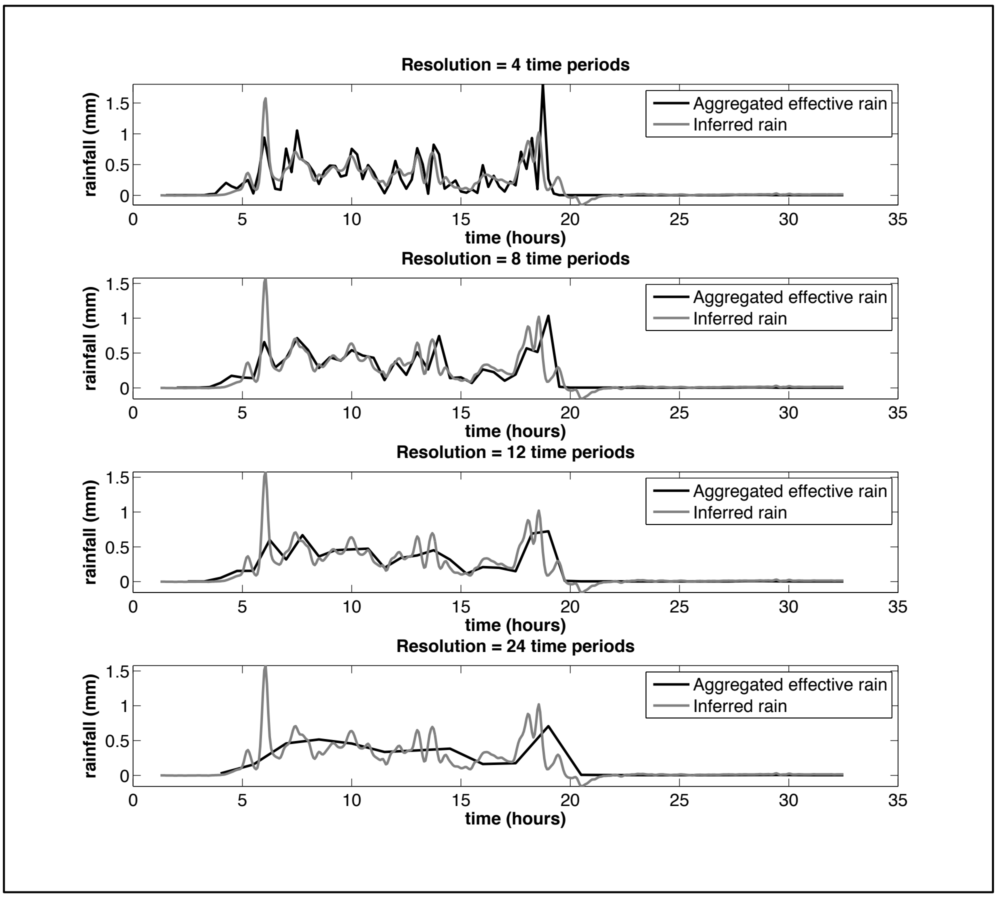
Periodograms of the amplitude spectra of the observed and modelled series were plotted to test whether the ER and IR have the same dynamics in the critical frequency range, despite the loss of time resolution (related to low pass filtering due to regularisation). A periodogram is the frequency domain representation of a signal; transforming the signal into the frequency domain may reveal information that is not visible in the time domain. A transfer function shown in its equivalent frequency domain form describes the mapping between the input and the output signals' spectra for the linear dynamic systems used here. Signals may be easily transformed between the time and frequency domains (Wickert, 2013).

Periodograms are obtained using the Matlab implementation via the Fast Fourier Transform and smoothed using the Integrated Random Walk (e.g. Young *et al*., 1999); the same regularisation approach as used in the calculation of the IR, implemented in the Captain Toolbox (Taylor *et al*., 2007). Periodograms of ER, IR and catchment streamflow are compared on a single plot showing how the spectral properties of the inversion process are used to obtain the IR estimates (see Fig. 6). The streamflow spectrum is the result of mapping the rainfall spectrum by the catchment dynamics. To make a full inversion of that mapping would involve very strong amplification of high frequencies with all the negative consequences discussed by Kretzschmar *et al.* (2014). The most significant implications of full inversion include the introduction of high amplitude, high frequency noise artefacts into the rainfall estimates. The regularisation of estimated derivatives introduces the effect of low-pass filtering into the inversion process, avoiding the excessive high frequency noise. Regularisation does not introduce any lag into the process, unlike traditional low pass filtering.

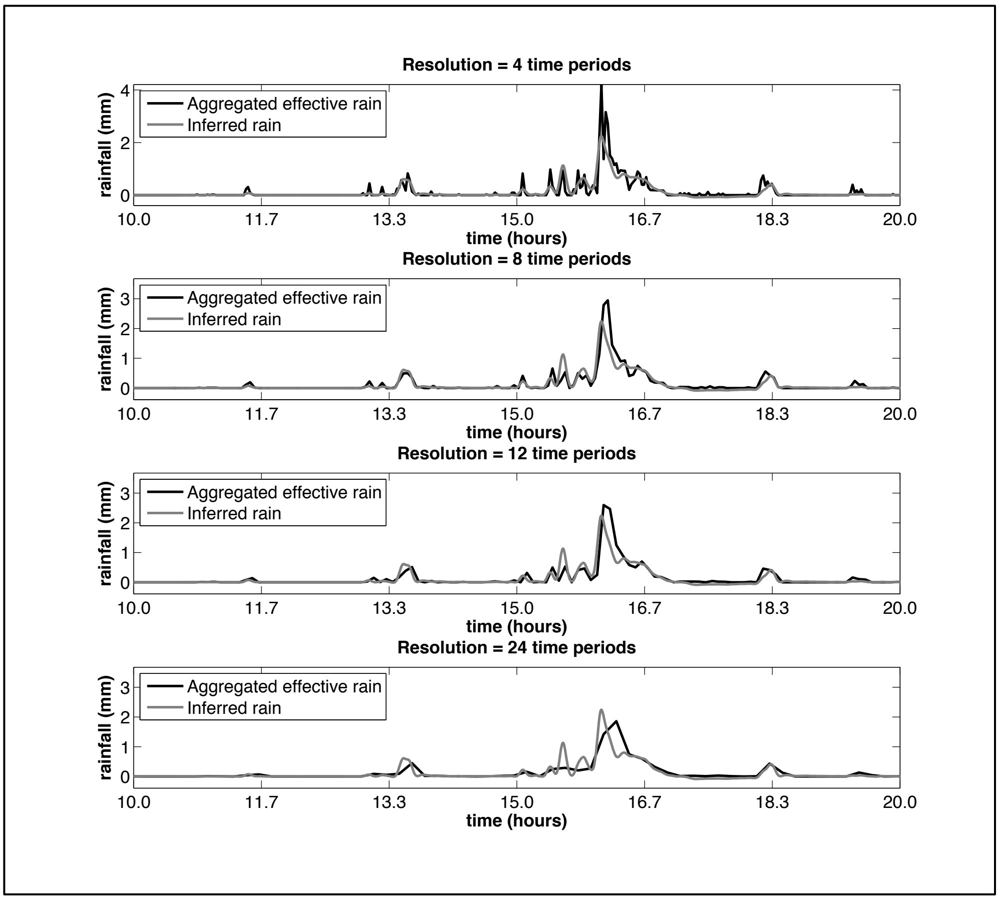
Results and discussion

Fig. 2 illustrates the smoothed rainfall distribution of the IR sequence obtained using the RegDer method. Similar streamflow sequences are generated using either the ER or IR sequences as model input (see Kretzschmar *et al*., 2014). The implication is that the catchment system dynamics are being captured despite the apparent difference in the rainfall distribution and that the detail of the rainfall series in time may not be important when modelling the dominant mode of streamflow dynamics.

In order to assess the degree of resolution lost by estimating rainfall using the RegDer method, the ER was aggregated using two methods (i.e. simple aggregation by resampling and a moving average) and the resulting sequences compared to the IR sequence in the time domain. Plots of progressively more aggregated sequences are shown in Fig. 3. It can be seen that as aggregation increases, peaks become lower and more spread out and the sequence is effectively smoothed. The coefficient of determination (R_t^2) and the correlation (R) between the aggregated sequence and the IR tends to a maximum then decreases as aggregation time increases – ultimately the variation in the sequence would be completely smoothed out. The point at which the maximum value is reached is taken as an estimate of the resolution of the IR. Plots of R_t^2 or R values are shown in Fig. 4 (aggregation by resampling) and Fig. 5 (moving average estimate). Time resolution estimates are shown in Table 1 and compared with the fast time constant (TC_q) and the Nyquist-Shannon sampling limit.

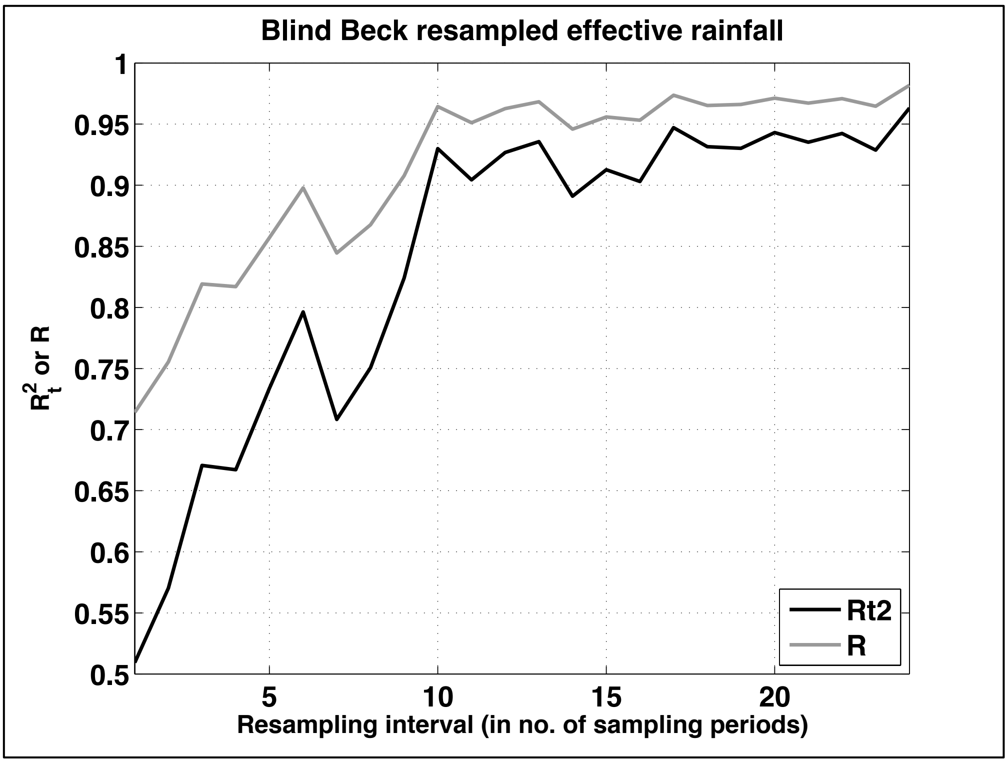

Resolution = 4 time periods
Resolution = 8 time periods
Resolution = 12 time periods
Resolution = 24 time periods
Aggregated effective rain
Inferred rain
rainfall (mm)
time (hours)

a)

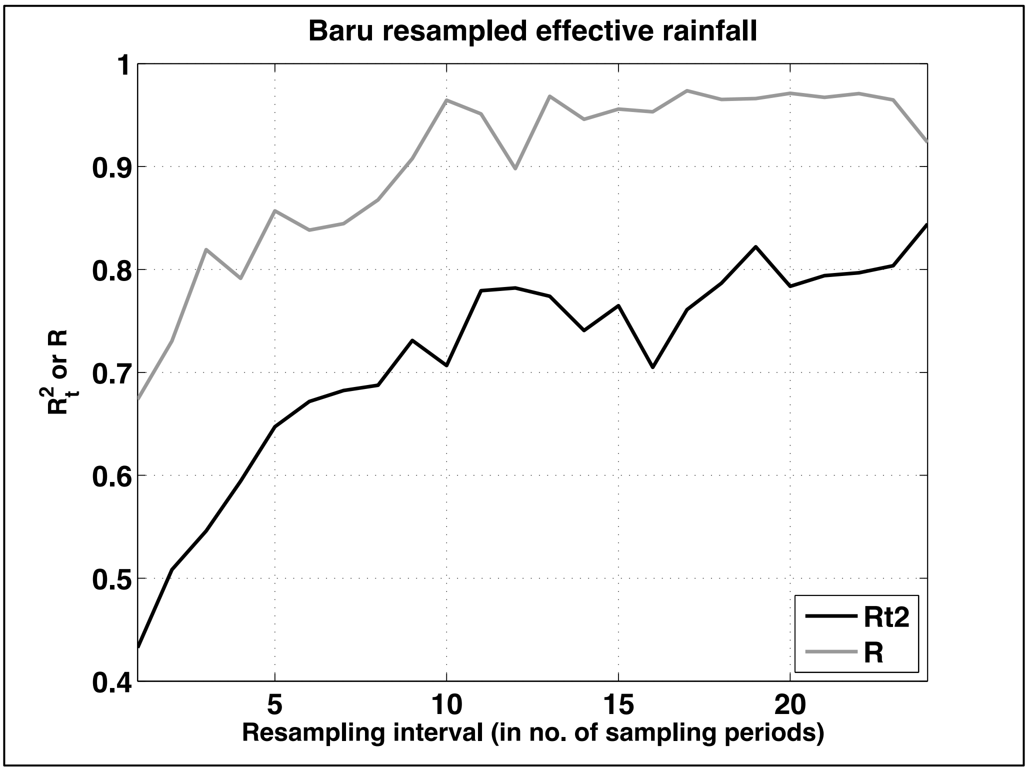

Resolution = 4 time periods
Resolution = 8 time periods
Resolution = 12 time periods
Resolution = 24 time periods
Aggregated effective rain
Inferred rain
rainfall (mm)
time (hours)

b)

Figure 3 – Comparison of aggregated sequence to the Inferred effective rainfall sequence for a)Blind Beck (sampling interval 15 mins) b) Baru (sampling interval 5 mins) at aggregations of 4, 8 12 and 24 time periods illustrating how aggregation lowers the peak and spreads the volume of rainfall over a longer time period. The inferred effective rainfall sequence is plotted for comparison.



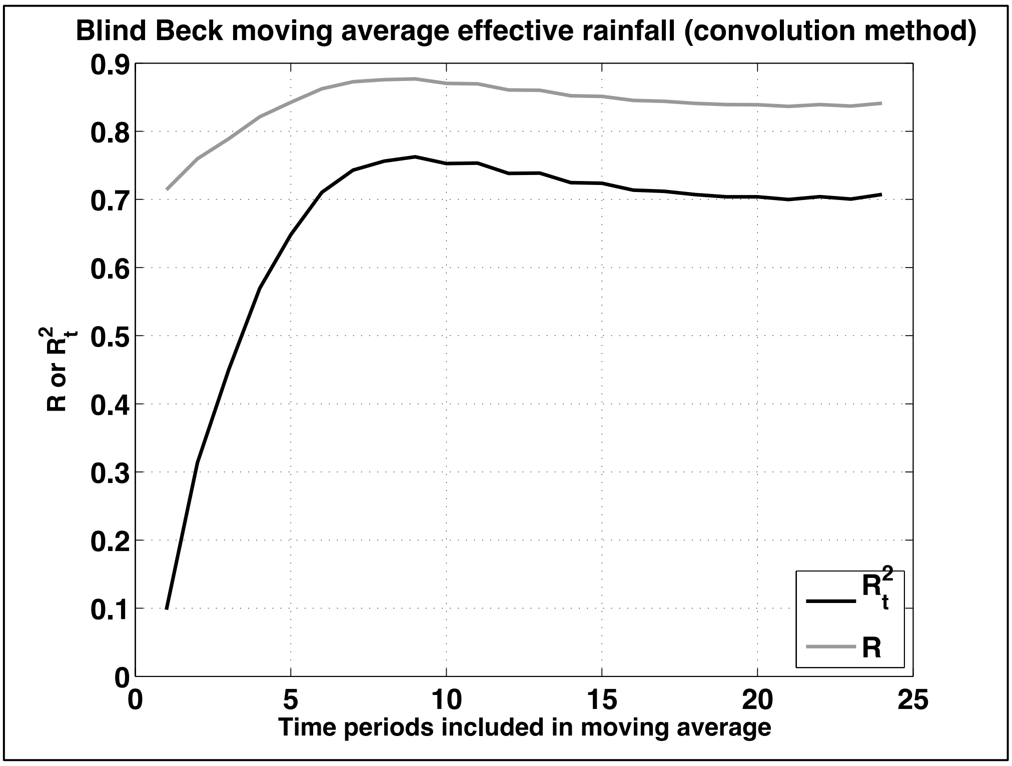
a)



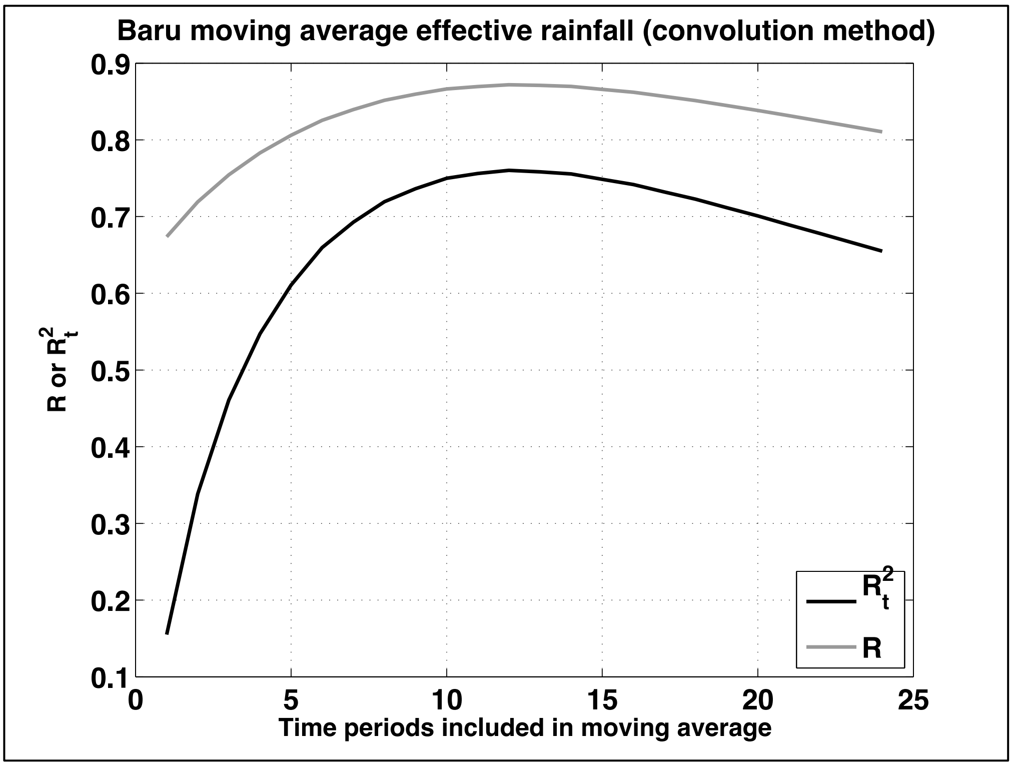
b)

Figure 4 – The R_t^2 and R tend to a maximum value as aggregation increases for a) Blind Beck and b) Baru. The resolution of the inferred effective rainfall is taken to be point at which the maximum is reached or very little change is apparent. For Blind Beck, this value is reached at 10 periods for both R_t^2 and R. The result for Baru is not quite as clear but can be estimated to be 10 periods from R and 11

or 12 from R_t^2 though R_t^2 continues to increase up to 24 time periods perhaps due to higher variability of the rainfall.



a)



b)

Figure 5 – A similar plot to Figure 4 with aggregation by Moving Average for a) Blind Beck and b) Baru. Rather than reaching an asymptotic level, the R_t^2 and R values maximize at 9 time periods for Blind Beck and 12 time periods for Baru (determined graphically in Matlab). These values have been used as the estimates of the resolution of the inferred effective rainfall and agree well with the estimates made by resampling.

Table 1 – Time resolution of the inferred effective rainfall sequences estimated by both resampling and moving average methods are less than the dominant (fast) mode of the catchments and considerably less than the 'safe' Nyquist-Shannon limit.

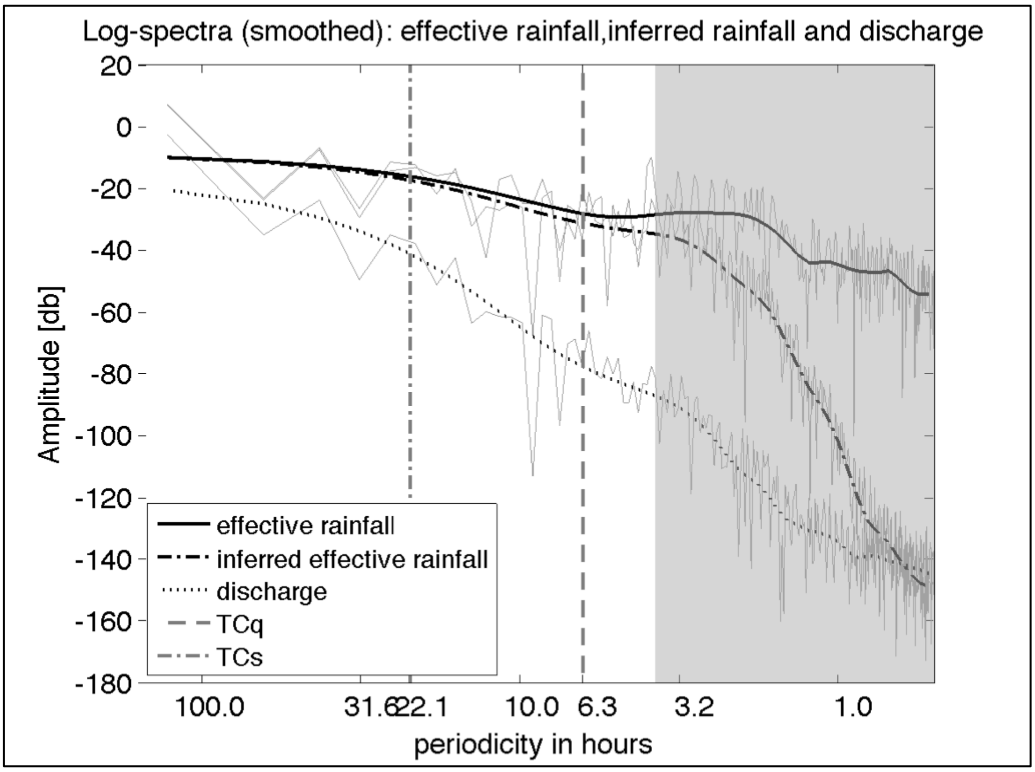
Catchment	Sampling frequency (hours)	TC_q (hrs)	TC_s (hrs)	SFI	Nyquist-Shannon Limit (hours)	Time resolution estimates: Aggregation by resampling	Time resolution estimates: Aggregation by Moving Average
Blind Beck	.25	6.3	22.1	47%	19.9	2.5 hours (10 time periods)	2.25 hours (9 time periods)
Baru	.083	1.1	18.7	62%	3.4	0.9 - 1 hours (11-12 time periods)	1 hour (12 time periods)

Table 1 shows that the estimated resolution of the IR sequence for Blind Beck is around 9-10 time periods (i.e. 2.25-2.5 hours) and for Baru it is 11-12 time periods (i.e. 55 mins – 1hr). Both estimates are within the Nyquist-Shannon safe sampling limit and below the fast time constant for both catchments indicating that even though resolution has been lost – the trade-off for numerical stability – the dominant mode of the rainfall-streamflow dynamics has been captured. Table 2 shows that the estimated resolution of the inferred effective rainfall for both catchments is well within the Nyquist limit and, whilst the Blind Beck resolution is within the safe limits suggested by Ljung (1999) and Young (2010), the estimated resolution for Baru is close to the fast TC and outside the suggested limits. The estimates of resolution of the inferred sequence made from the aggregation plots are not always well-defined and may be dependent on the length of record which will affect the number of aggregation periods that may be meaningfully calculated given the finite length of the data series. A better means of estimation of resolution may be achieved by examining the frequency spectra of the rainfall and streamflow sequences.

Table 2 – The estimated resolution of the inferred effective rainfall for Blind Beck is well within both the Nyquist limit and the safe sampling limits suggested by the Ljung (1999) and Young (2010) whereas the resolution Baru, whilst well within the

Nyquist limit, is close to the fast TCq and outside the suggested safe sampling limits of Ljung and Young.

Catchment	TCq (hours)	Nyquist limit (hours)	Ljung interval (hours)	Young interval (hours)	Estimated resolution (hours)
Blind Beck	6.3	19.9	3.98	3.32	2.25-2.5
Baru	1.1	3.4	0.68	0.57	0.91-1.0

a) 

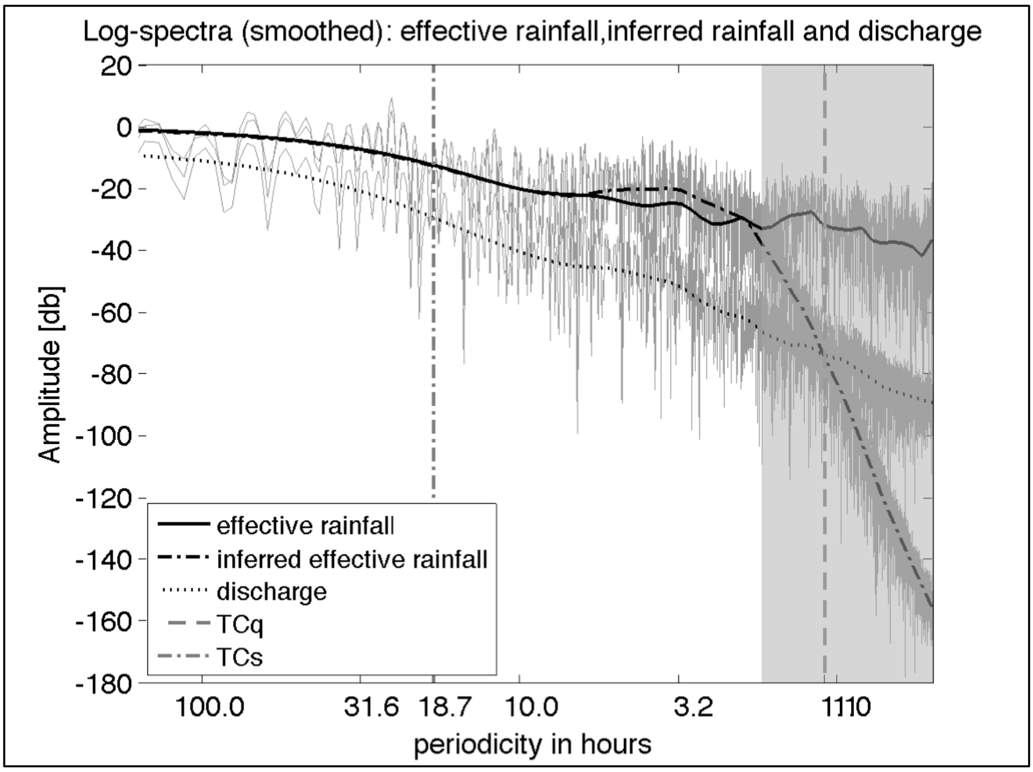
b) 

Figure 6 – Periodograms for a) Blind Beck and b) Baru showing the frequency structure of the effective rain, inferred effective rain and streamflow sequences. Both catchments show a similarity in the frequency spectra of effective and inferred effective rainfall within the catchment system. The inferred effective rainfall spectrum is very close to the actual effective rainfall one within a wide range of frequencies mostly covering those corresponding to the catchment's time constants. There is also a strong low pass filtering effect cutting off high frequencies with low amplitudes instead of boosting this high frequency noise.

In Figure 6, the amplitude spectra of inferred effective and observed effective rainfall are very close (overlapping when smoothed) within a broad range of frequencies. The cut-off frequency, where the difference between the smoothed ER and IR spectra is approximately -6Db, provides a frequency domain estimate of the resolution. The cut-off period for Blind Beck is 3.8 hours and for Baru is 1.7 hours. For frequencies above this value, a very strong low pass filtering effect shown is by the rapid decrease in the IR spectrum. The frequency range beyond the cut-off point, shaded in Fig. 6, carries a very small proportion of the power of the signal and can be considered non-significant.

The processes and characteristics limiting the inferred effective rainfall accuracy include the slow components of the catchment dynamics and the rainfall regime. These can be seen as the ‘usual suspects’ affecting the inversion process. The general goodness of fit of the initial catchment model (rainfall-streamflow) appears to be a factor as well, indicating that the inferred effective rainfall estimation method presented here can be used to assess the quality of available data and the degree to which the data characterise the catchment.

Conclusions

A combination of time and frequency domain techniques have been used to show that the inferred effective rainfall time-series generated by the RegDer inversion method does indeed approximate the direct inverse of a transfer function to a high degree of accuracy within the frequency range which includes the dominant modes of the rainfall-streamflow dynamics. The direct inverse exaggerates low-amplitude high frequency noise, which is filtered out by the regularisation process involved in the RegDer method. The smoothing of the signal resulting from regularisation is quantified in the time-domain by comparison with aggregated observed input data using standard model fit measures - coefficient of determination, R_t^2, and correlation coefficient, R - and analysed as a low-pass filtering process in the frequency domain.

Acknowledgements

The authors would like to thank Mary Ockenden for the collection and quality assurance of the period of rainfall and streamflow for the Blind Beck catchment (NERC grant number NER/S/A/2006/14326), and also Jamal Mohd Hanapi and Johnny Larenus for the collection of the period of rainfall and streamflow utilised for the Baru catchment and to Paul McKenna for its quality assurance (NERC grant number GR3/9439). This work has been partly supported by the Natural Environment Research Council [Consortium on Risk in the Environment:

Diagnostics, Integration, Benchmarking, Learning and Elicitation (CREDIBLE)] grant number: NE/J017299/1.

References

Andrews, F., Croke, B., Jeanes, K., 2010 Robust estimation of the total unit hydrograph. In: *2010 International Congress on Environmental Modelling and Software Modelling for Environment's Sake. Ottawa, Canada.*

Beven, K.J., 2006 A manifesto for the equifinality thesis. *J. Hydrol*. 320, 18-36.

Beven, K.J., 2012 *Rainfall-Runoff Modelling - the Primer*, second ed. John Wiley and Sons, Chichester, England.

Bidin, K., Chappell, N., 2003 First evidence of a structured and dynamic spatial pattern of rainfall within a small humid tropical catchment. *Hydrol. Earth Syst. Sci*. 7 (2), 245-253.

Bidin, K., Chappell, N., 2006 Characteristics of rain events at an inland locality in Northeastern Borneo, Malaysia. *Hydrol. Process*. 20 (18), 3835-3850.

Bloomfield, P. 1976 *Fourier Analysis of Time Series: An Introduction*, John Wiley & Sons, New York.

Blöschl, G. (Editor), Sivapalan, M. (Editor), Wagener, T. (Editor), Viglione, A. (Editor), Savenije, H. (Editor), 2013 *Runoff prediction in ungauged basins: synthesis across processes, places and scales*, Cambridge University Press

Boorman, D. 1989 A new approach to Unit Hydrograph Modelling. *PhD Lancaster*

Brocca, L., Moramarco, T., Melone, F., & Wagner, W. 2013 A new method for rainfall estimation through soil moisture observations. *Geophysical Research Letters*, 40(5), 853-858.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., Levizzani, V., 2014 Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research: Atmospheres*, 119(9), 5128-5141.

Chapman T. G. 1996 Common unitgraphs for sets of runoff events: Part 2. Comparisons and inferences for rainfall loss models. *Hydrological Processes*, **10**:783–792

Chappell, N., Franks, S., Larenus, J., 1998 Multi-scale permeability estimation for a tropical catchment. *Hydrol. Process*. 12 (9), 1507-1523.

Chappell, N.A., Tych, W., Chotai, A., Bidin, K., Sinun, W., Chiew, T.H., 2006 Barumodel: combined data based mechanistic models of runoff response in a managed rainforest catchment. *For. Ecol. Manag*. 224 (1), 58-80.

Croke, B., 2006 A technique for deriving an average event unit hydrograph from streamflow-only data for ephemeral quick-flow-dominant catchments. *Adv. Water Resour*. 29 (4), 493-502.

Clark, M. P., and D. Kavetski 2010 Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes, *Water Resour. Res.,* 46, W10510, doi:10.1029/2009WR008894

Cuchi, J., Chinarro, D., Villarroel, J., Antonio Cuchi, D., & Luis Villarroel, J. 2014 Linear system techniques applied to the Fuenmayor Karst Spring, Huesca (Spain). *Environmental Earth Sciences*, 71(3), 1049-1060.

Duband, D., Obled, C., Rodriguez, J.-Y., 1993, Unit hydrograph revisited: an alternative approach to UH and effective precipitation identification, *Journal of Hydrology*, 150, 115-149

Holman-Dodds, J. K., Bradley, A. A. and Sturdevant-Rees, P. L., 1999 Effect of temporal sampling of precipitation on hydrologic model calibration, *Journal of Geophysical Research*, Vol. 104, No. D16, Pages 19,645-19,654

Jakeman, A.J., Littlewood, I.G., Whitehead, P.G., 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology* 117, 275-300.

Kavetski, D., and M. P. Clark 2010 Ancient numerical daemons of conceptual hydrological modelling: 2. Impact of time stepping schemes on model analysis and prediction, *Water Resour. Res*., 46, W10511, doi:10.1029/2009WR008896

Kirchner, J., 2009 Catchments as simple dynamical systems: catchment characterization, rainfall-runoff modelling, and doing hydrology backward. *Water Resour. Res*. 45.

Krajewski, W. F., Lakshmi, V., Georgakakos, K. P., & Jain, S. C. 1991 A Monte Carlo study of rainfall sampling effect on a distributed catchment model. *Water Resources Research*, 27 (1), 119-128.

Kretzschmar, A., Tych, W and Chappell, N. A. 2014 Reversing hydrology: Estimation of sub-hourly rainfall time-series from streamflow. *Environmental Modelling & Software* 60: 290-301.

Krier, R., Matgen, P., Goergen, K., Pfister, L., Hoffmann, L., Kirchner, J.W., Uhlenbrook, S., Savenije, H.H.G., 2012 Inferring catchment precipitation by doing hydrology backward: a test in 24 small and mesoscale catchments in Luxembourg. *Water Resour. Res*. 48, W10525.

Littlewood, I. G. 2007 Rainfall–streamflow models for ungauged basins: uncertainty due to modelling time step. In: *Uncertainties in the 'Monitoring–Conceptualisation– Modelling' Sequence of Catchment Research (Proc. Eleventh Biennial Conference of the Euromediterranean Network of Experimental and Representative Basins)* (L. Pfister & L. Hoffmann, eds), 149–155. Paris: UNESCO Tech. Doc. in Hydrology Series 81.

Littlewood, I.G. and Croke, B.F.W. 2008 Data time-step dependency of conceptual rainfall—streamflow model parameters: an empirical study with implications for regionalisation, *Hydrological Sciences Journal*, 53:4, 685-695, DOI: 10.1623/hysj.53.4.685

Littlewood, I. G., Young, P. C. and Croke, B. F. W. 2010 Preliminary comparison of two methods for identifying rainfall–streamflow model parameters insensitive to data time-step: the Wye at Cefn Brwyn, Plynlimon, Wales. In: *Proceedings of the Third International Symposium (British Hydrological Society, 19–23 July 2010, Newcastle University, UK).*

Littlewood, I.G. and Croke, B.F.W. 2013 Effects of data time-step on the accuracy of calibrated rainfall-streamflow model parameters: practical aspects of uncertainty reduction, *Hydrology Research*, 44.3

Ljung, L. 1999. *System identification. Theory for the user*. Prentice Hall, Upper Saddle River, 2nd edition.

Mayes, W.M., Walsh, C.L., Bathurst, J.C., Kilsby, C.G., Quinn, R.F., Wilkinson, M.E., Daugherty, A.J., Connell, P.E., 2006 Monitoring a flood event in a densely instrumented catchment, the Upper Eden, Cumbria, UK. *Water Environ. J.* 20 (4), 217-226.

Montanari, A., and E. Toth 2007 Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?, *Water Resour. Res*., 43, W05434, doi:10.1029/2006WR005184.

Ockenden, M.C., Chappell, N.A., 2011 Identification of the dominant runoff pathways from data-based mechanistic modelling of nested catchments in temperate

UK. *J. Hydrol*. 402 (1), 71-79.

Ockenden, M.C., Chappell, N.A., Neal, C., 2014 Quantifying the differential contributions of deep groundwater to streamflow in nested basins, using both water quality characteristics and water balance. *Hydrol. Res*. 45, 200-212.

Olivera, F., Maidment, D., 1999 Geographical information systems (GIS)- based spatially distributed model for runoff routing, *Water Resources Research*, 35, 1155-1164

Reynard, N.S., Stewart, E.J., 1993 The derivation of design rainfall profiles for upland areas of the UK. *Meteorol. Mag*. 122, 116-123.

Szolgayová, E.,; Arlt, J., Blöschl, G., Szolgay, J. 2014 Wavelet based deseasonalization for modelling and forecasting of daily streamflow series considering long range dependence. *Journal of Hydrology and Hydromechanics*, 62 (1). pp. 24-32. ISSN 0042-790X

Taylor, C.J., Pedregal, D.J., Young, P.C., Tych, W., 2007 Environmental time series analysis and forecasting with the Captain toolbox. *Environ. Model. Softw*. 22 (6), 797-814.

Teuling, A.J., Lehner, I., Kirchner, J.W., Seneviratne, S.I., 2010 Catchments as simple dynamical systems: experience from a Swiss pre-alpine catchment. *Water Resour. Res.* 46 (10).

Wickert, M., 2013 *Signals and Systems for Dummies*, 1st Edition, John Wiley & Sons;

Young, P.C. 2010 The estimation of continuous-time rainfall–flow models for flood risk management. In*: Proceedings of the Third International Symposium (British Hydrological Society, 19–23 July 2010, Newcastle University, UK).*

Young, P.C. and Beven, K.J., 1994 Data-based mechanistic modelling and the rainfall-flow non-linearity, *Environmetrics*, 5, 335-363

Young, P.C., Garnier, H., 2006 Identfication and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems. *Environ. Model. Softw*. 21 (8), 1055-1072.

Young, P.C., Sumisławska, M.A., 2012. Control systems approach to input estimation with hydrological applications. In: *16th IFAC Symposium on System Identification. July 2012, Brussels, Belgium, pp. 1043-1048*.

Young, P.C., Pedregal, D.J., Tych, W., 1999 Dynamic harmonic regression. *J. Forecast.* 18 (6), 369-394.