Journal of Hydrology 634 (2024) 131034

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Convolution, deconvolution, the unit hydrograph and flood routing

John D. Fenton

Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Karlsplatz 13/222, 1040 Vienna, Austria

ARTICLE INFO

This manuscript was handled by Sally Elizabeth Thompson, Editor-in-Chief, with the assistance of Jose-Luis Molina, Associate Editor.

Keywords:
Rainfall
Runoff
Transfer function
Inflow
Outflow
Simulation

ABSTRACT

Convolution equations are used to relate the input and the output of a system such as rainfall and runoff, or inflow and outflow of a river reach. There have been numerous reports of unsatisfactory results from the deconvolution necessary to calculate the connecting transfer function. The cause is that the equations are ill-conditioned, and it is shown here that the fundamental theoretical solution is that of wild oscillations such as has often been found computationally. A spectral method is proposed for numerical solution, where, instead of individual point values, the transfer function is expressed as series of given continuous functions, where the problem is to determine the coefficients of those functions. The resulting equations have been found to be well-conditioned, and solutions obtained were smooth, bounded, and enabled a certain amount of physical interpretation of the transfer function. The method has been applied to several problems, including typical rainfall–runoff ones and flood routing and wave propagation problems, with quite satisfactory results. Another problem for deconvolution is found to be the traditional use of truncated equations. A remedy is only to use later output data points where convolution with input data does not reach back beyond the initial one. For the routing of larger flood events, the linear methods employed were found to be not so accurate. However as they are a first approximation that requires no knowledge of stream geometry or resistance, and as either discharge or water level hydrographs can be used, they may be useful.

1. Introduction

In many rainfall–runoff and flood routing problems little is known about the physical nature of the catchment, the stream or streams, their geometry or their resistance. However, if one knows input and output time series, systems techniques can be used to model even a complicated river system. One of the pillars of elementary hydrology has been the use of such linear systems theory, with an input, usually rainfall, and an output, river flow at a station, the two connected by a unit hydrograph, such that the output is expressed as a convolution, a weighted sum of preceding values of the input. Here throughout we use the term "unit hydrograph" not in the sense where a finite period of input is considered, but a single spike of input, elsewhere known as the "*instantaneous* unit hydrograph". With this, variations of rainfall during the course of the storm can be considered without some of the laborious considerations for a finite duration. Also here we do not concern ourselves with details of preliminary operations such as subtraction of base flow.

The fundamental assumptions are those of time invariance, where the parameters of the model do not change with time, and linearity such that all contributions can be combined additively. Dooge (1973, p18) wrote "in hydrology, the assumptions of linearity and time-invariance are not valid, but nevertheless have been used for a long time in applied hydrology because of the simplification they introduce".

The first problem in an application is usually that of deconvolution, determining the sequence of values of the transfer function, from given input and output data. After that it can then be used with other inputs to determine the corresponding outputs. The numerical solution, deconvolution, has always been found to be difficult. The problem of actually obtaining a unit hydrograph from data has been, to quote Hamlet, "more honoured in the breach than the observance". It has often been mentioned, but relatively rarely addressed. When it has, the results have usually been alluded to darkly as being not so satisfactory. Even the immense theoretical work of Dooge (1973) contained no numerical results and little mention of difficulties. The problem has been blamed on a variety of causes. For example, Chow et al. (1988, p218) wrote "the resulting unit hydrograph may show erratic variations and even have negative values ... [this] may be due to nonlinearity in the effective rainfall–direct runoff relationship in the watershed, and even if this relationship is truly linear, the observed data may not adequately reflect this". Zhao et al. (1995) considered more possibilities: "... one often encounters a situation in which the derived unit hydrograph exhibits noise fluctuation among its ordinates. This could

E-mail address: johndfenton@gmail.com.
URL: http://johndfenton.com/.

https://doi.org/10.1016/j.jhydrol.2024.131034
Received 4 October 2023; Received in revised form 26 November 2023; Accepted 3 December 2023
Available online 11 March 2024

be caused by nonlinearity, time-variance, distributed-space, sampling intervals for time, numerical instability in computation, measurement errors, and multicollinearity". Other elementary books on hydrology declare themselves in favour of the method, yet rarely is it actually described how deconvolution should be performed, or problems alluded to, or results presented. An exception is Shaw (1994).

There have been attempts to overcome the difficulties by modifications of the formulations. For example, in the "Ridge least-squares method" a positive-valued number, called the ridge parameter, is added to the diagonal elements of the matrix obtained from a least-squares formulation of the convolution equations (Zhao, Tung, and Yang, 1995). Another approach is that of Yang and Han (2006), transforming the convolution using Z transforms and then in the resulting complex plane restricting the nature of singularities.

Problems of deconvolution in the civil engineering application of the mixing characteristics of urban drainage structures were addressed by Stovin, Guymer, Chappell, and Hattersley (2010) using Regularisation, in which a single objective function is constructed under a series of constraints encapsulating a priori knowledge regarding the function to be estimated. In subsequent works from the same school, most recently Sonnenwald, Mark, Stovin, and Guymer (2021), similar sophisticated techniques were used, including maximum entropy deconvolution. In other fields where the general problem of deconvolution has been considered, there exist powerful signal processing software and works that have used them. Press, Teukolsky, Vetterling, and Flannery (1992, §13) contains a full discussion plus programs. Deconvolution by Fourier transform is mentioned as being almost trivial, but the vulnerability is mentioned if a Fourier component of the input is small, which can easily happen, as we shall see. There are also warnings: "... the process of deconvolution has other practical shortcomings ... generally quite sensitive to noise in the input data, and to the accuracy to which the response function is known. Perfectly reasonable attempts at deconvolution can sometimes produce nonsense for these reasons. In such cases you may want to make use of the additional process of optimal filtering ...". Such an approach is used in the routine TFESTIMATE in the mathematical software package MATLAB, which involves sectioning the record and averaging modified periodograms of the sections.

Two current hydrologic software packages can be mentioned; both avoid systematic deconvolution. The Hydrologic Modeling System of the US Army Corps of Engineers Hydrologic Engineering Center (HEC-HMS, 2023), uses synthetic hydrographs and only mentions deconvolution when the reader is referred to Chow, Maidment, and Mays (1988). The Revitalised Flood Hydrograph model of the UK Centre for Ecology and Hydrology (Kjeldsen, 2007) similarly just mentions Chow et al. (1988), and uses a simple kinked triangular-shaped unit hydrograph involving a time-to-peak parameter, catchment area and the selected time step.

The unit hydrograph concept is a strong simplification of the complexity of basin hydrology. Cudennec (2005, p221) notes that "its globality does not easily allow accounting for space heterogeneities and variabilities; it applies only to rapid components of runoff; its linearity and stationarity can be criticised because actual hydrologic events are nonlinear; and the claim of describing all transfer processes within a basin is strong regarding the differences between hillslope and channel processes". That has not stopped the application of linear systems methods as a first approximation.

For flood routing, where both input and output are measured hydrographs, the concepts underlying the unit hydrograph are rather more applicable. Sauer (1971, 1973) seems to be the first to have actually implemented the idea of using convolution for river reaches. He cited an earlier paper by Dooge and Harley, who wrote "it is remarkable that a general linear analysis of the type used in unit hydrograph procedures has not been applied also to the problem of channel routing". In Sauer's work, there is no mention of deconvolution. The transfer functions were obtained by taking a single inflow spike input and routing it down the stream by rather approximate methods. His work was complemented by that of Keefer and McQuivey (1974) and Keefer (1974, 1976), and the development of a software package by the US Geological Survey, described in Doyle, Shearman, Stiltner, and Krug (1984). That body of work, while using the highly approximate determination of the transfer function, then included attempts at nonlinear generalisation and greater accuracy by including different travel times and behaviour for different magnitudes of input.

Since that body of work some 40–50 years ago, there has been little mention of the methods and their use in flood routing. Goring (1984) seems to have been the first in the context of river routing to have performed deconvolution — and by using Fourier methods, which make this simpler, however in hydrology there is an extensive literature on the problems of high-frequency oscillations associated with that approach.

Now to set the background to the present paper and to describe what it does. The problems associated with deconvolution seem to have had a powerful dissuasive effect on its use. In this work it is suggested that the most fundamental and devastating reason for that is the ill-conditioning of the equations, when any one equation is very much like the next. It is shown how a theoretical solution of such equations is a wildly-oscillating one, which is what the previous numerical works have discovered. Then, conventional analysis methods are described, including least-squares solution of the convolution equations and Fourier deconvolution, which have the difficulties described above. An alternative, a family of spectral methods is suggested here, based on approximating the transfer function by series of simple mathematical functions (Fourier, polynomial, and cubic splines), suppressing the tendency to wild oscillation. When applied, the methods all agreed closely. For practical problems, good solutions with short series suffice. If, however, longer series are used, they contain the seeds of their own destruction, for they "over-fit" the data and tend to the oscillatory behaviour noted for point-wise methods.

Once the problems of deconvolution have been solved, the use of transfer functions seems to be robust and flexible. The methods are also applied here to flood routing, where the input is a hydrograph, the water level (stage) or discharge at one station as a function of time, and the output is one at a downstream station. Stage hydrographs are usually directly and easily measured, and can be used here without the need to use rating curves at either station to work in terms of discharge. In any case the water level is often the most important quantity.

For the above-mentioned assumption of time-invariance to apply there can be no variable control such as gates or valves at the downstream station. Otherwise, with a fixed control such as a weir, or channel control where the stream is flowing freely without impediment, any backwater effect is included in the underlying physics of the problem. However, the second major assumption, that of linearity, is only satisfied in an approximate sense, albeit rather better than in rainfall–runoff problems.

2. The deconvolution problem

Consider the discrete convolution formulation

$$O_n = \sum_{k=0}^{K} h_k \, I_{n-k}, \quad \text{for} \quad n = 0, 1, \dots, \tag{1}$$

giving values of the output sequence O_n in terms of a weighted sum of the current and preceding values of the input sequence I_n, I_{n-1}, …. The weights are the h_k where k denotes the "reach-back" time level, constituting the (discrete) transfer function, a finite sequence of values, $\left(h_k : k = 0, \dots, K\right)$. It is assumed that both input and output sequences commence at I_0 and O_0. However in the first K terms of the series in Eq. (1), we need information from I_{-K} to I_{-1} which we do not have. There are two common solutions to that problem:

- One is to assume that all those values are zero, so that we modify Eq. (1) so as to write, using two different but equivalent notations

$$O_n = \sum_{k=0,k\leqslant n}^{K} h_k\, I_{n-k} = \sum_{k=0}^{\min(n,K)} h_k\, I_{n-k}, \quad \text{for } n = 0, 1, \dots, \tag{2}$$

the second having as upper limit the minimum of the two numbers n and K. One does not evaluate contributions I_{n-k} for $k > n$, with the implication that all contributions I_m for $m < 0$ are zero. So as not to have a finite discontinuity between -1 and 0, throughout calculations in the present work all input and output sequences had the initial values subtracted such that $I_0 = O_0 = 0$ also. This seems to be common practice, where the base flow must be separated from the stream flow hydrograph in the derivation of a unit hydrograph and must subsequently be added to the flow derived using unit hydrograph techniques to obtain total flow.
- Alternatively, if one uses Fourier methods, one can assume that the I_m, the O_n, and the h_k are all periodic, with the period that of the data sets N so that each data value for $m < 0$ is the same as that at the corresponding other end of the data $m \pm N$. That is a special assumption that is necessary for the convenience of Fourier deconvolution, as will be seen below, but which has its effect on results.

Consider the nature of the system of Eqs. (1) or (2) with three general adjacent terms

$$O_n = \cdots h_{n-m} I_m + h_{n-m+1} I_{m-1} + h_{n-m+2} I_{m-2} + \cdots, \tag{3}$$

and considering that and the next equation for O_{n+1}, and writing as a matrix equation:

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & I_m & I_{m-1} & I_{m-2} & \cdots \\ \cdots & I_{m+1} & I_m & I_{m-1} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \cdots \\ h_{n-m} \\ h_{n-m+1} \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots \\ O_n \\ O_{n+1} \\ \cdots \end{bmatrix}. \tag{4}$$

The two matrix rows are made up of two identical sequences, displaced by one element, but where the coefficients, the input I values, vary smoothly and slowly, so that the rows are so similar to each other, that, well-known for such a case, the matrix is ill-conditioned and solutions can contain large fluctuations.

A simple solution demonstrating such behaviour can be obtained. Consider just three convolution equations (2) and the three first contributions:

$$O_n = h_0 I_n + h_1 I_{n-1} + h_2 I_{n-2} + \cdots \tag{5a}$$

$$O_{n+1} = h_0 I_{n+1} + h_1 I_n + h_2 I_{n-1} + \cdots \tag{5b}$$

$$O_{n+2} = h_0 I_{n+2} + h_1 I_{n+1} + h_2 I_n + \cdots. \tag{5c}$$

As a first approximation in the spirit of a Taylor series, we assume that locally the inflow and outflow are varying linearly in time and so they do also in the corresponding sequences, and we write $I_{n+i} = a_0 + i\, a_1 + \cdots$ and $O_{n+i} = b_0 + i\, b_1 + \cdots$ for $i = -2, \dots, +2$, and substitute into the three Eqs. (5). We find just two independent equations, with h_2 indeterminate, and the surprisingly revealing solutions:

$$h_0 = +h_2 + \text{terms in } a \text{ and } b \text{ coefficients}, \tag{6a}$$

$$h_1 = -2h_2 + \text{terms in } a \text{ and } b \text{ coefficients, and} \tag{6b}$$

$$h_2 = +h_2, \tag{6c}$$

showing how the solution for the transfer function seeks to oscillate wildly in the sense +1, −2, +1 shown — even where the input and output sequences are varying smoothly, or not at all!

Generally, there are more points at which the convolution is evaluated than there are terms in the transfer function, and the system of Eqs. (2) is over-determined, which is just as well, as some equations such as the first, $I_0 h_0 = O_0$ cannot be solved explicitly, given our assumptions that I_0 and O_0 are zero. However solving as part of an over-determined system it is not such a problem. We note that if we had not subtracted to zero, the first equation would give the solution $h_0 = O_0/I_0$, which is a strong and simple statement requiring that finite value for h_0 just to give initial agreement between the two sequences. Whatever that h_0 is, in view of the oscillatory solution in Eqs. (6), it can be anticipated that h_1 is a finite opposite-signed value, which is typical of what one finds in computations.

In view of the artificiality of the augmentation of zero values $I_m = 0$ for $m < 0$ expressed as the truncation of the convolution equations Eq. (1) to those of Eq. (2), it will be found not always desirable to consider all convolutions from $n = 0$ containing the non-physical values of zero inflow. We generalise the notation to consider the convolution equations for points in the interval $n = [n_{\min}, n_{\max}]$.

Similarly we introduce a generalisation for the domain of k, giving the length of the transfer function. The classical convolution starts at $k = 0$, meaning that the output sequence is computed from values of the input sequence commencing at that very moment. In problems of flood routing that we want to consider, it may take some time for the effect of a change of input to be experienced as output, as the body of a flood moves downstream. We will consider the possibility of a non-zero minimum value $k_{\min}$. Introducing $k_{\max}$ for the maximum value, the domain of k (the extent of possible values) is the interval $\left[k_{\min}, k_{\max}\right]$.

The general expression of the convolution equations can then be written

$$O_n = \sum_{k=k_{\min}}^{\min(n,k_{\max})} h_k\, I_{n-k}, \quad \text{for } n = n_{\min}, \dots, n_{\max}. \tag{7}$$

The conventional expression used throughout the literature is to use this with *all* output data points, thus setting $n_{\min} = 0$ (and using $k_{\min} = 0$). This means that the first convolution equation, for example, is $O_0 = h_0 I_0$, with the implication that one has assumed $I_m = 0$ for $m < 0$ and this clearly-inadequate equation, and subsequent ones successively involving more terms, is taken as playing a role in determining h_0, h_1, and so on. It will be demonstrated in examples in the results Section 5 that this can give poor results. The remedy is only to use convolution equations starting at a later point that involve only known values of input.

3. Existing solution methods

3.1. Least-squares solution of the point-wise convolution equations

Consider the convolution equations (7) with the general domains in n and k written in matrix form

$$\mathrm{I}\,\mathbf{h} = \mathbf{O} \tag{8}$$

where I is the $(n_{\max} - n_{\min} + 1) \times (k_{\max} - k_{\min} + 1)$ matrix with elements I_{n-k}, $\mathbf{h}$ is a column vector with $k_{\max} - k_{\min} + 1$ elements h_k, and $\mathbf{O}$ a column vector with $n_{\max} - n_{\min} + 1$ elements O_n.

The usual procedure for deconvolution has been to obtain a least-squares solution.of Eq. (8). Press et al. (1992, §15.4) present the method and program for a Singular Value Decomposition solution that can be used. Alternatively one can solve the over-determined system using optimisation methods, minimising the sum of squares of the errors in all the convolution equations. The common method that can be used is that described, for example, in Chow et al. (1988, §7.6) and Shaw (1994, §13.4.2). The procedure is to pre-multiply both sides by the transpose I^{T}, giving

$$\mathrm{I}^{\mathsf{T}}\mathrm{I}\,\mathbf{h} = \mathrm{I}^{\mathsf{T}}\mathbf{O}, \tag{9}$$

giving a square $(k_{\max} - k_{\min} + 1) \times (k_{\max} - k_{\min} + 1)$ matrix pre-multiplying the vector of unknowns, and on the right a $(k_{\max} - k_{\min} + 1)$ vector. The equation can be solved by common software. However, the ill-conditioning of the coefficient matrix often gives irregular highly oscillatory solutions for the transfer function, referred to throughout the literature, and shown in Section 2 above here.

3.2. Fourier deconvolution

One method of deconvolution using Discrete Fourier Transforms (DFT) was introduced by O'Donnell in 1960. Because Fourier series are orthogonal under summation, taking the DFT of a convolution gives a simple product in Fourier coefficient space, so that transforming input and output gives an algebraic expression for the Fourier coefficients of the transfer function, which can then be inverted to give the values of the function itself. Naturally, the linearity of the system means that no new frequency components are generated. In principle, the method is simple and powerful. In application it is quite vulnerable. The theory is given here, as some notational matters are novel to this field.

As Fourier series are being used, the implication is that each of the quantities, the input and output sequences and the transfer function, are all of equal length N, numbered 0 to $N-1$, and are periodic. First we define the DFT and its inverse.

Discrete Fourier transform. Consider a sequence f_n, for $n = 0, 1, \dots, N-1$. The transform of such a sequence is denoted by $F_j = \mathcal{D}(f, j)$:

$$\begin{aligned} F_j &= \frac{1}{N}\sum_{n=0}^{N-1} f_n \exp(-\mathrm{i}2\pi jn/N) \\ &= \frac{1}{N}\sum_{n=0}^{N-1} f_n W^{-jn}, \text{ for } j = -N/2, \dots, +N/2, \end{aligned} \tag{10}$$

where $W = \exp(\mathrm{i}2\pi/N)$, the fundamental Nth root of unity, and where $\mathrm{i} = \sqrt{-1}$. This form contains two conventions: first, the transform contains the factor $1/N$ such that the Fourier coefficients F_j have a similar magnitude to the f_n (for example, F_0 is actually the mean of the f_n sequence). Secondly, this version has minus signs in the exponential terms. The reason will shortly be explained.

Inverse discrete Fourier transform. This is defined to be $f_n = \mathcal{D}^{-1}(F, n)$:

$$f_n = \sum_{j=-N/2}^{N/2}{}'' F_j \exp(+\mathrm{i}2\pi jn/N) = \sum_{j=-N/2}^{N/2}{}'' F_j W^{+jn}, \text{ for } n = 0, \dots, N-1, \tag{11}$$

where the range of j is the symmetric one, $j = -N/2, \dots, +N/2$ rather than from 0 to $N-1$. The notation Σ'' means that contributions to the sum at the ends, $j = \pm N/2$, are to be multiplied by $1/2$. The positive exponent in this inverse, Eq. (11), is such that the interpolating Fourier series of the f_n sequence, using t as the continuous variable, with a period T, also contains a positive sign

$$f(t) = \sum_{j=-N/2}^{N/2}{}'' F_j \exp(+\mathrm{i}2\pi jt/T), \tag{12}$$

as are the series for the derivatives of this interpolating function without any extra minus signs. A happy circumstance is that if the series in Eq. (12) is evaluated from $j = -J, \dots, +J$, where some $J < N/2$, it is the corresponding least squares approximation to the f_n sequence.

Convolution theorem for periodic sequences. Now returning to convolution matters, consider where the I, O, and h sequences are all periodic, each with N terms, $0, \dots, N-1$. In this case the general convolution Eq. (1) can be simplified, not by truncation as was done to give Eq. (2), but one can now include terms I_{n-k} for $n-k < 0$ and it can be written simply

$$O_n = \sum_{k=0}^{N-1} h_k I_{n-k}, \quad \text{for} \quad n = 0, 1, \dots, N-1. \tag{13}$$

Taking the DFT of that convolution, from Eq. (10),

$$\mathcal{D}(O, j) = \frac{1}{N}\sum_{n=0}^{N-1}\left(\sum_{k=0}^{N-1} h_k I_{n-k}\right) W^{-jn} = \sum_{k=0}^{N-1} h_k \left(\frac{1}{N}\sum_{n=0}^{N-1} I_{n-k} W^{-jn}\right), \tag{14}$$

where the order of summation has been changed. Multiplying by a term W^{-jk} inside the first (outer) summation and by the inverse of that W^{+jk} inside the next (inner) summation,

$$\mathcal{D}(O, j) = \underbrace{\sum_{k=0}^{N-1} h_k W^{-jk}}_{N\,\mathcal{D}(h,j)} \underbrace{\left(\frac{1}{N}\sum_{n=0}^{N-1} I_{n-k} W^{-j(n-k)}\right)}_{\mathcal{D}(I,j)}, \tag{15}$$

where the inner sum of a function of $n-k$ over n just means taking all the N values irrespective of k because of periodicity, so it is simply the DFT of I. The result is that the DFT of the convolution is a simple product of Fourier terms

$$\mathcal{D}(O, j) = N\,\mathcal{D}(h, j)\,\mathcal{D}(I, j). \tag{16}$$

This gives an explicit solution for the jth term of the transform of the transfer function in terms of transforms of input and output:

$$\mathcal{D}(h, j) = H_j = \frac{1}{N}\frac{\mathcal{D}(O, j)}{\mathcal{D}(I, j)}. \tag{17}$$

The $1/N$ factor has entered here because of the inclusion of $1/N$ in the present definition of the transform, Eq. (10). Goring (1984) used a different definition without the extra factor.

The transfer function itself is then given by the inverse DFT $h_k = \mathcal{D}^{-1}(\mathcal{D}(h, j), k)$. It can now be used with other input sequences to determine the corresponding outputs. Of course, the DFTs and inverses can be conveniently obtained using Fast Fourier Transform software.

In practice the behaviour of solutions can be very erratic. Considering the denominator of Eq. (17), $\mathcal{D}(I, j)$ for $j = -N/2, \dots, +N/2$, small values can be encountered, somewhat randomly, giving large values of the H_j with unreasonably large contributions at that frequency, and highly oscillatory solutions for the h_k. As values of $\mathcal{D}(I, j)$ depend crucially on the data, any irregularities can lead to greatly different results. In the present work, for the more idealised examples presented further below, it was found that the solutions were sufficiently oscillatory that this method of Fourier deconvolution could not be used. However, Goring (1984) obtained satisfactory solutions for several Aotearoa New Zealand rivers, noting that the Fourier spectrum decayed sufficiently quickly that one could ignore components above a certain frequency.

3.3. Solving convolution equations by optimisation

A recent paper by the author Fenton (2023) proposed a different deconvolution method, formulating the solution of the convolution equations as an optimisation problem. A generalisation was that there could be a number of different inputs, the general convolution Eq. (7) written as a sum over J different tributary contributions and here written as equation e_n, the value of which the solution method should try to minimise:

$$e_n = O_n - \sum_{j=1}^{J} \sum_{k=k_{\min}}^{\min(n, k_{\max,j})} h_{j,k} I_{j,n-k}, \quad \text{for } n = n_{\min}, \dots, n_{\max}. \tag{18}$$

where $h_{j,k}$, $I_{j,n-k}$, and $k_{\max,j}$ are simple generalisations of the h_k, I_{n-k}, and $k_{\max}$ used above. The programming and solution of such a system, such that the sum of all the e_n^2 is minimised, was relatively simple using optimisation software. This was used in a study of flows in a complex set of interconnections in the Broken River valley in south-eastern Australia. This was from an era and situation where readings were only taken daily. The transfer functions were then all very short, typically just three values, corresponding to zero, one and two days delay before upstream input appeared as output at the downstream station. The method worked well for the problem described, where questions of oscillatory long-term transfer functions did not occur.

The author also applied the optimisation solution method to some of the single-input problems described below in Section 5. He discovered, to his naive surprise, that the results were exactly the same as those of the least-squares method for solution of the convolution equations described above in Section 3.1, with the attendant problems of oscillatory transfer functions. It was the realisation of that which led the author then to develop the family of methods described in the next section.

However, the optimisation approach described here could still have certain applications where transfer functions are short-run. A possibility is that it would allow the use of nonlinear formulations. Setting up of a problem is relatively simple, and optimisation software robust.

4. Deconvolution methods using approximation by spectral methods

Here, a different approach to deconvolution will be considered. Rather than sophisticated signal-processing methods being applied to the original formulation, the problem is recast using spectral methods. Instead of the transfer function h_k being composed of a number of different point values, each obtained from solution of the "point-wise" equations as described above, the values of h_k are approximated by a series of given continuous functions of the index k and where the problem now is to determine the coefficients of those functions by least-squares methods, where the approximating series is used with integer k in convolution expressions. The subsequent success of this approach was because it is one of approximation with low-order relatively smooth functions, not attempting to describe the natural tendency to oscillation discovered in Section 2. Three families of functions of k are considered.

4.1. Approximating functions

4.1.1. Fourier series

The functions considered here were almost the same as those used for Fourier convolution in Section 3.2, but here using shorter series and not using complex notation. For our purposes, to be able to approximate any finite variation of $h(k)$ the form of the Fourier series is not quite complete. The constant $j = 0$ term, F_0 in Eq. (12), which is the mean value of the f_n sequence in that case, needs to be augmented by a linear term. This corresponds to the well-known procedure in Fourier analysis of subtraction of a linear trend line from a signal, so that the implied periodicity does not cause a discontinuity between the beginning and end of the signal. We introduce a scaled k-like variable $\theta_k = 2\pi\left(k - k_{\min}\right) / \left(k_{\max} - k_{\min}\right)$ varying from 0 to 2π as k goes from $k_{\min}$ to $k_{\max}$ so that we write

$$h_k = h(k) = a_0 + b_0\theta_k + \sum_{m=1}^{M}\left(a_m \cos m\theta_k + b_m \sin m\theta_k\right), \tag{19}$$

The problem is to determine the a_m and b_m for $m = 0, 1, 2, \ldots$.

4.1.2. Polynomial series

To represent the $h(k)$ an obvious alternative possibility would be simply a polynomial made up of monomial terms such as

$$h_k = h(k) = \sum_{m=0}^{M} c_m\left(k - k_{\min}\right)^m, \tag{20}$$

where c_m for $m = 0, \ldots, M$ are the coefficients to be determined, M is the degree of the polynomial, and we represent variation with k as the monomial quantity $\left(k - k_{\min}\right)^m$. There are two problems with this. As k is an integer quantity, $\left(k - k_{\min}\right)^m$ can become very large, the coefficients c_m would have to become very small and the solution process might be more demanding. The second difficulty is that the set of monomials all look the same (consider, starting at $x = 0$, the monomials x^2, x^3, x^4, …). The author, Fenton (2018), in a study of the rating curves at measurement stations for discharge as a function of stage, considered such polynomial approximation at some length, theoretically and computationally. He concluded that if one has to approximate data with arbitrary variation, it is better to use series of Chebyshev polynomials. In a similar manner to terms in Fourier series, each has different behaviour, enabling a more efficient form of approximation and with the ability to represent more general non-periodic behaviour.

Consider the series of Chebyshev polynomials of the first kind T_m (see, for example, Abramowitz and Stegun, 1965, §22) written

$$h_k = h(k) = \sum_{m=0}^{M} c_m T_m\left(\kappa_k\right), \tag{21}$$

where κ_k is the scaled k-like quantity

$$\kappa_k = 2\frac{k - k_{\min}}{k_{\max} - k_{\min}} - 1, \tag{22}$$

such that the domain of κ_k (the range of validity) is from -1 to $+1$ as k varies from $k_{\min}$ to $k_{\max}$. In fact (Abramowitz and Stegun, 1965, eqn 22.3.15) the Chebyshev polynomials are simply evaluated and thought of as

$$T_m\left(\kappa_k\right) = \cos\left(m \arccos \kappa_k\right). \tag{23}$$

The behaviour of the polynomials can be seen in Abramowitz and Stegun (1965, fig 22.6), or Fenton (2018, fig 2) where comparison with monomials is made. For small m they look like simple polynomials: $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$; for increasing m they oscillate ever more rapidly between ± 1, resembling higher terms in Fourier series, with the same ability as them to approximate general behaviour, but without any implied periodicity.

4.1.3. Approximating splines

The Fourier and Chebyshev polynomial series are both *global* approximations, valid over the whole interval. Problems of such are well-known: if behaviour anywhere is too rapid, there is a general loss of accuracy. Given the possible irregular behaviour of transfer functions, $h(k)$ as a function of k, an alternative is to use piecewise continuous approximation in the form of cubic splines such as applied by Fenton (2018) as a second method for describing rating curves. The domain of k is divided into a small finite number of sub-intervals, and a different cubic polynomial applies in each of them. The $h(k)$ are to be approximated over each interval by a polynomial of low degree $M = 2$, 3, or more. So, over each of the intervals j, h is approximated as a function of a continuous k by a different function

$$h_j(k) = \sum_{m=0}^{M} c_{j,m}\left(k - k_j\right)^m. \tag{24}$$

where k_j is the value at the beginning of interval j. Certain continuity conditions are applied such that the function and derivatives up to $M-1$ agree on both sides of the internal knot points separating the intervals.

4.2. Choice and implementation of polynomial approximation

Each of the formulations, Fourier series, polynomial series, and approximating splines, were tested in detail using all the example problems presented below in Section 5. It was found that all worked well. The three formulations agreed sufficiently closely, that it is considered necessary here only to present details and results for one and recommend that for use: the polynomial series are the simplest to present, and gave the best results, by a narrow margin.

In general we evaluate the convolutions from $n = n_{\min}$ to $n = n_{\max}$; we may not want to evaluate a convolution which requires truncation at $n = 0$, involving artificial zeroes. The system of Eqs. (7) becomes, substituting Eq. (21) and interchanging orders of summation

$$O_n = \sum_{m=0}^{M} c_m\left(\sum_{k=k_{\min}}^{\min(n,k_{\max})} T_m\left(\kappa_k\right) I_{n-k}\right), \quad \text{for } n = n_{\min}, \ldots, n_{\max}. \tag{25}$$

The problem is now to solve for the c_m, $m = 0, \ldots, M$. The system can be written

$$\mathrm{C}\,\mathbf{c} = \mathbf{O}, \tag{26}$$

where C is a $(n_{\max} - n_{\min} + 1) \times (M + 1)$ matrix, with elements

$$C_{nm} = \sum_{k=k_{\min}}^{\min(n,k_{\max})} T_m(\kappa_k)\, I_{n-k}, \tag{27}$$

where n runs from $n_{\min}$ to $n_{\max}$ and m from 0 to M (taking liberties here with conventional matrix element numbering); $\mathbf{c}$ is a column vector with elements c_m, $m = 0, \dots, M$ and $\mathbf{O}$ a column vector with elements O_n, $n = n_{\min}, \dots, n_{\max}$. Whereas in the Fourier series the orthogonality under summation of the exponential functions meant that it was possible to obtain an explicit solution, the Chebyshev polynomials have different orthogonality properties, as shown for the continuous forms in Abramowitz and Stegun (1965, §§22.1&2). Here no such simplicity seems possible, or even desirable in view of the behaviour of the explicit Fourier coefficients. What is most useful is that the polynomials are all different, enabling more general representation. For solution it is necessary to use linear algebra software, as throughout all the approximation methods here.

Eq. (26) can be solved in a least-squares sense by Singular Value Decomposition, such as the program of Press et al. (1992, §15.4). More common software for square matrix problems can be implemented by pre-multiplying both sides of Eq. (26) by the transpose $C^{\intercal}$, similar to that done to give Eq. (9), with the result

$$C^{\intercal} C\, \mathbf{c} = C^{\intercal} \mathbf{O}, \tag{28}$$

with a relatively small well-conditioned square $(M + 1)\times(M + 1)$ matrix pre-multiplying the $M+1$ vector of unknowns, and on the right an $M+1$ vector.

In the testing described below with finite small values of $M = 6$ to 16, the matrix C did not suffer from ill-conditioning; results for the transfer function $h(k)$ were smooth, accurate, and showed some physical significance (such as an obvious maximum corresponding to the separation distance between peaks of inflow and outflow). For larger M values however, the solution became more oscillatory. As the series contained more rapidly-oscillating functions the method began to look more like the traditional point-wise convolution form. This is evidence of what is called *over-fitting* in statistics, where an approximating method is allowed too much freedom to agree with data. It seems that the satisfactory smooth results obtained for low levels of approximation are to be preferred, but a certain amount of judgement is necessary.

5. Examples

In the various examples to be considered, all four of the main methods mentioned above were applied, using the point-wise convolution equations in Section 3.1 plus the three spectral formulations, Section 4.1.

5.1. A rainfall–runoff problem

Shaw (1994, figure 13.11) gave a typical example of the problems presented in elementary hydrology textbooks with rainfall and corresponding runoff. For this work the figure was digitised, with 106 equally-spaced values for rainfall and runoff with time interval $\Delta t = 0.45\,\mathrm{h}$ and is shown in Fig. 1. The figure reveals a feature of the present approach, that we are trying to develop methods that can handle any data sequence. The rainfall data in blocks of constant duration in the original, suited to methods that purport to handle input data of a given duration, is here represented as a continuous sequence of values at intervals rather smaller than the block duration. The values are constant in a particular block, but with the methods advocated here, considering only instantaneous unit hydrographs, they could have any variation.

The various deconvolution methods described above were applied, using a transfer function with $n_{\min} = 0$ and the maximum length of the transfer function $k_{\max}\Delta t \approx 30\,\mathrm{h}$, and with the levels of numerical approximation shown in part (b) of the figure. It can be seen in part (a) that all four methods, including the traditional point-wise approach described in Section 3.1, gave results for the computed outflow, obtained as part of the solution from the computed transfer function, which seem quite satisfactory. In fact, an attempt was made to separate the data into two, corresponding to the two main rainfall events, to give one set for determination of the unit hydrograph and another for validation, to test the results, however there was too much overlap between the events and no sensible results were obtained. Part (b) of the figure shows the main results of interest here, those of the transfer function h_k, now using different line types. The point-wise approach has led to finite oscillations, while the spectral methods gave smooth functions that could be used confidently with other input data.

5.2. A more complex rainfall–runoff problem with an important result

Kjeldsen (2007, figure 6.5) presented data for a "notable event" peaking on 26 January 2000 at Lennoxlove, east of Edinburgh in Scotland. The digitised data with 95 points in time are shown here in Fig. 2; the rainfall amounts, measured hourly, are quite irregular. Unlike as was done here for Fig. 1 each block was represented by a single rainfall value at its centre.

The figure also shows the computational results obtained here with a total transfer function length of $k_{\max}\Delta t = 2\,\mathrm{d}$. Just results using polynomial approximation are shown. Performing the calculations revealed something important that the simpler problem of Fig. 1 did not. Initially the traditional approach was taken with $n_{\min} = 0$ (and $k_{\min} = 0$), meaning that from Eq. (2), the first equation used is $O_0 = h_0 I_0$ and so on. It is well-known that the first points of any conventional convolution are unreliable, as these truncated short-run contributions contain insufficient information with the implication of the artificial zeroes $I_m = 0$ for $m < 0$. Fig. 2(a) with the highly variable input values shows that the results for the computed outflow as part of the solution, shown by small points, were unexpectedly poor: highly irregular and often far from the actual outflow data.

If one then assumes that only convolution equations should be included that contain actual input values I_m, $m \geqslant 0$, then the solution process should not start before a point $n_{\min} = k_{\max}$ such that the initial convolution equation does not reach back before 0. For example, taking the usual case $k_{\min} = 0$, such that the convolution includes input contributions right up to the point considered, the first convolution equation is then $O_{n_{\min}} = h_{k_{\max}} I_0 + \cdots + h_0 I_{k_{\max}}$. More generally for non-zero $k_{\min}$, the convolution equations are then, substituting $n_{\min} = k_{\max}$ into Eq. (7), the minimum value of $(n, k_{\max})$ becomes just $k_{\max}$, giving:

$$O_n = \sum_{k=k_{\min}}^{k_{\max}} h_k\, I_{n-k}, \quad \text{for } n = k_{\max}, \dots, n_{\max}, \tag{29}$$

such that we consider all input data values I_m from $m = 0$ but no equations or output data values for $n < n_{\min} = k_{\max}$.

The results using this are shown by the solid lines on Fig. 2(a), where using the spectral approach (with $M = 16$ for both polynomial and Fourier methods, and with 16 intervals for splines) good agreement with data was obtained. It does not matter that we do not plot results for times such that $n < n_{\min} = k_{\max}$, we are not predicting, we are merely showing results that are part of the solution process, to be able to evaluate its accuracy.

The usefulness of what has been done is reflected in part (b) of the figure, showing how the truncated convolutions, starting at $n = 0$ gave very different results for the transfer function from those using the procedure suggested here using convolution equations only for points $n \geqslant k_{\max}$.

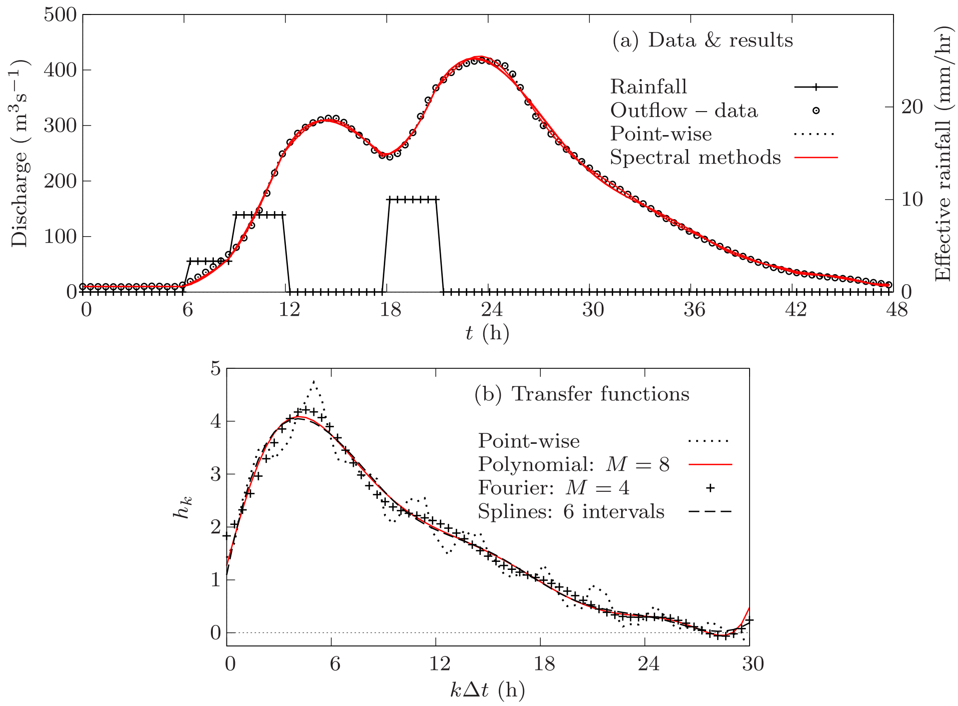

Fig. 1. A rainfall–runoff problem (Shaw, 1994, figure 13.11).

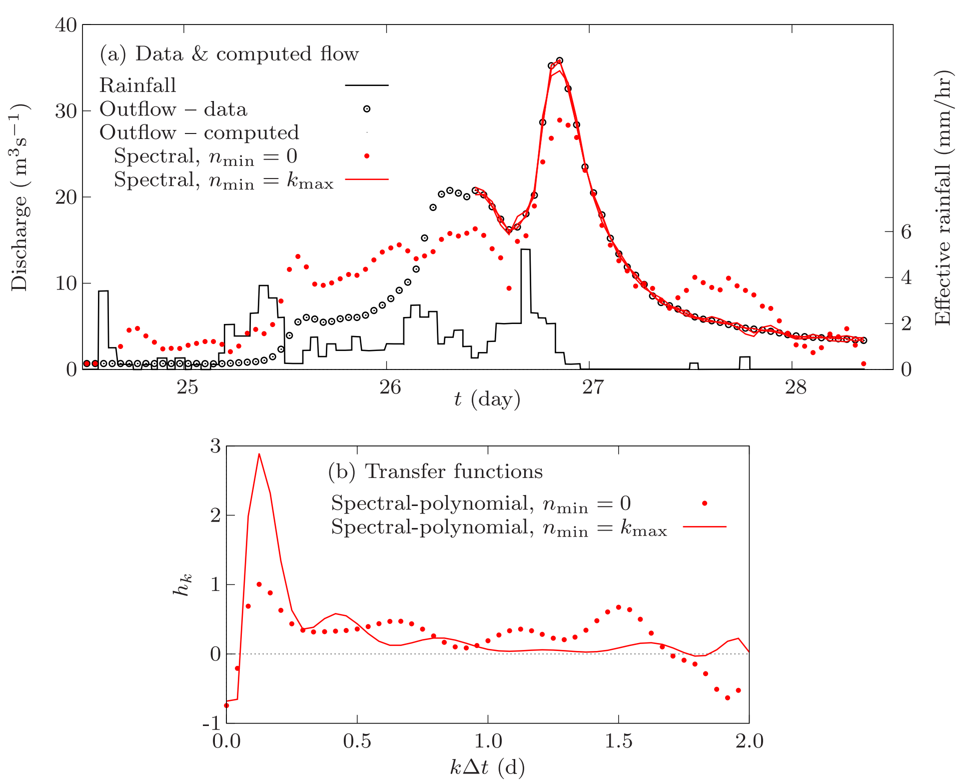

Fig. 2. Irregular rainfall data with a storm hydrograph.
Source: Taken from Kjeldsen (2007, figure 6.5).

5.3. A flood routing problem

The next example considered is a flood routing one, with flow at two stations on the Wye River, Erwood in Wales and Belmont in England, given in Flood Studies Report (1975, figure 3.6). The figure was digitised, with 74 equally-spaced interpolated values for inflow and outflow with time interval $\Delta t = 1.6\,\mathrm{h}$. The various methods described above were applied. The total length of the transfer function was 2.5 d. For the spectral methods, parameters were $M = 8$ for polynomials, $M = 4$ for Fourier, with 6 spline intervals. Results are shown in Fig. 3. Part (a) shows the overall problem with data plus computed results obtained as part of the system identification. All methods – point-wise and the

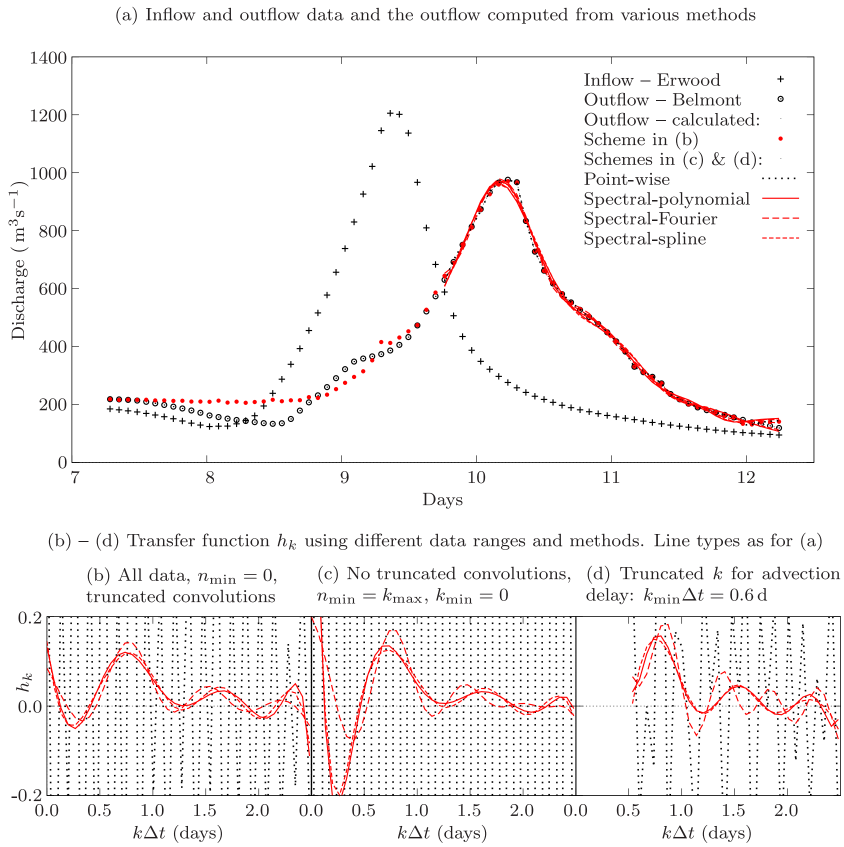

Fig. 3. A flood routing problem on the River Wye.

three spectral ones – give very similar results for the outflow computed as part of the solution. The thickened line is actually made up of six, three different spectral methods for two different values of k_{min}. Results were relatively insensitive to the number of terms in the approximating functions. What was more important for system identification was what points were considered for convolution and what was the length of the transfer function. There were three different cases considered, with the results for h_k being shown in figure parts (b) to (d).

Scheme in Fig. 3(b), $n_{min} = 0$, $k_{min} = 0$. This is the conventional method using all convolution equations from $n = 0$. In Section 5.2 with an irregular input it was found to give poor results. Here with smoother input the problems are fewer, but Fig. 3(a) shows how it still takes some time for a convolution to be accurate, to incorporate all the values in the transfer function — initially the solution is horizontal due to the dominance of the phantom zero data points before $n = 0$. Gradually agreement with measured outflow is better until after point $n = k_{max}$ at a time of about 9.8 d, agreement with outflow data and all other computational results, obscured by lines from other results, is very good.

The important result for this work is, however, contained in the next part (b) of the figure, showing the computed transfer functions. The traditional point-wise least squares approach (Section 3.1) gave wild oscillations of h_k between large positive and negative values, as foreseen in the mathematical solution of Eqs. (6). The spectral approach adopted here (Section 4) has performed well, with smooth bounded variation, such that the results could confidently be combined with other input flood hydrographs to compute outputs. The results show a maximum in h_k at about 0.7 d, corresponding to the separation of the flood peaks in part (a), and with lesser maxima that we can interpret as harmonics of the dominant wave. What is less easily explicable is the increase of h_k as $k \to 0$, but we have already seen above that the behaviour of the transfer function is not obvious, such as when we showed the fundamental oscillating solution in Eqs. (6).

Scheme in Fig. 3(c), $n_{min} = k_{max}$, $k_{min} = 0$. The next part shows results obtained by only evaluating convolution equations where the whole transfer function plays a role, $n \geq k_{max}$, such that every convolution point uses all the terms in the transfer function, as was done in Section 5.2. Oscillations of h_k in (c) are rather larger than in part (b). The reason is that for the first convolution point at 9.8 d the inflow shows a massive flood peak which has passed at that moment and the flow is decreasing quickly, but that is yet to be felt downstream so that the downstream outflow is increasing. By using the conventional $k_{min} = 0$ in our convolutions here, allowing outflow to be immediately partly determined by inflow, we are saying that the outflow at 9.8 d is being partly given by the large upstream event, of which, of course, it actually still knows nothing. The contortions of the transfer function required to handle this irrationality can be seen in the larger fluctuations of h_k for smaller k. What is possibly surprising, however, is that the transfer function can still handle this problem at all, and give the good agreement with outflow data shown by the solid lines for $t > 9.8$ d in part (a). However, the behaviour shown in part (c) as the anomalous large oscillation for small $k\Delta t$ would have possibly deleterious consequences if the transfer function were convolved with other data. The oscillations of the results from the point-wise method were even wilder in this case where, obviously, they had to work harder.

Scheme in Fig. 3(d), $n_{min} = k_{max}$, $k_{min}\Delta t = 0.6$ d. In flood routing problems usually there is a finite time for input to appear as an output. It is expected that the first terms of the transfer function will be zero, so one can set the time of the first arrival $k_{min}\Delta t$ to be a finite number.

Fig. 3(d) shows the results for h_k obtained by inspecting the data in (a) and choosing $k_{min}\Delta t = 0.6$ d, so that, subtracting from the first convolution point at 9.8 d on the figure, we have, possibly arbitrarily, suggested that nothing of the inflow after about 9.2 d has yet reached the downstream end. Part (d), with this adjustment, now shows the most physical results of all three h_k figures — there is no sudden

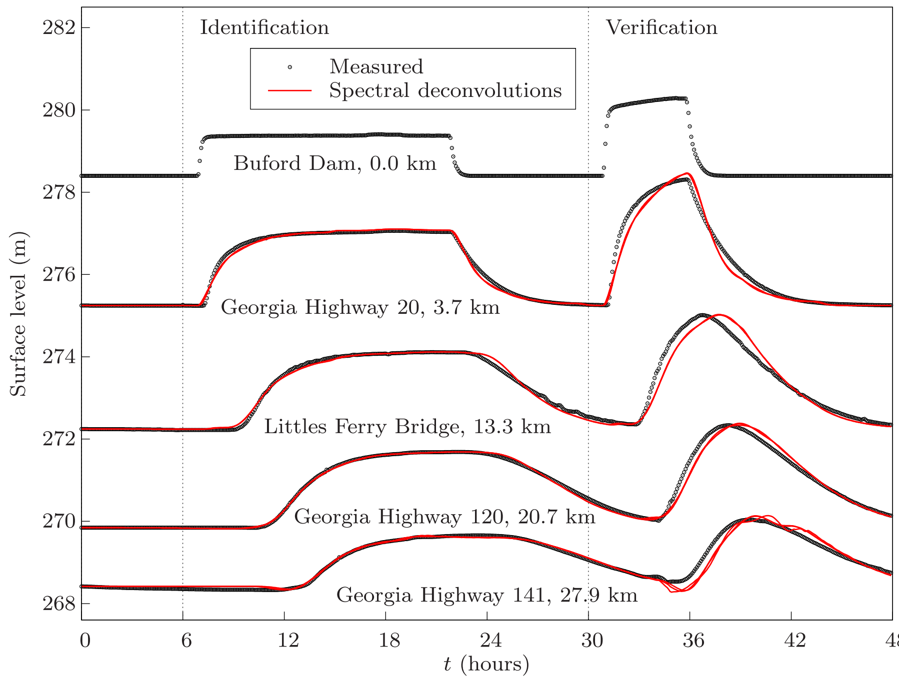

Fig. 4. Simulation of waves in four reaches on the Chattahoochee River from two sudden flow releases, showing the identification range for $6 \leq t \leq 30$ h, and the testing range thereafter.

increase for small k, and the oscillations of the figure seem to have a physical basis, that of the delay of the main wave plus harmonics. Correspondingly, the point-wise results also were less wild in this case.

One last observation however, in view of the three cases studied, is that the transfer function h_k, provided it can be computed reliably, is a robust means of representation. Throughout, the approach seems to have worked well despite the difficulties mentioned.

5.4. Comparison with field measurements of hydraulic transients

Faye and Cherry (1980) describe a large-scale experiment performed on a 28 km reach of the Chattahoochee River in the vicinity of Atlanta, Georgia, USA. Collection of stage and discharge data occurred intensively over a 3-day period in 1976. Initial flow conditions in the river were steady and low. Commencing early in the morning of 22 March the discharge at the upstream point, a hydro-electric dam, was increased and maintained at that rate for about 20 h before returning to low flow. On 23 March, a second controlled discharge was released, resembling a typical hydro-power wave or pulse. Measurements of stage at a number of stations were made continuously, and measurements of flow were also made, which are not needed here, and neither are other measurements such as the physical dimensions of the river.

The stage measurements are ideal for testing the methods in this work. Fig. 4 shows five hydrographs. Here data from 6 – 30 h were used to determine the transfer functions, of lengths 8 h, for four successive reaches. Then, convolutions were used to obtain the stage hydrographs also for the following 18 h, our verification or testing phase. It was found here that the traditional truncated convolutions gave the best results. This might be due to the abrupt almost-discontinuous nature of the data. The results plotted are those from all deconvolution methods: point-wise, and the three spectral methods. The results are for the cascade of individual reach combinations, 1–2, 2–3, 3–4, and 4–5 and not, for example, 1–4. All four methods show close agreement, including the traditional point-wise approach. However the transfer functions from the point-wise method were highly oscillatory. Agreement is generally good, however there is a consistent tendency for the predicted wave to arrive after the actual one, to travel more slowly. This may be partly due to the shorter length of the test wave or its magnitude. The question arises: how large were the surges? From the data for hydraulic radius given in Table 2 of Faye and Cherry (1980) a typical original depth of approximately 1.5 m is found, so that the increase in depth of the first slug of about 1.8 m and the second of about 2.5 m were rather more than the original depth, so they were quite large, those for the testing phases slightly larger, and hence faster, than those of the identification phases. They are also shorter, which might have played a role. In the next section we consider some more demanding tests of the effects of nonlinearity.

5.5. Testing physical accuracy limits

The underlying theory throughout has assumed that the physical system is a linear one. Yet the long wave equations describing the propagation of flood waves are nonlinear, including effects such that higher waves travel faster. The equations also show diffusion and dispersion (when modelled by linear theory!), where different wavelength components travel at different speeds and show different rates of diminution. The present approach is expected to be able to describe that, however.

We consider two different tests, for each of which two simulations were conducted, each time introducing an input hydrograph into a uniform channel, and accurately numerically solving the unsteady nonlinear long wave equations by an implicit finite difference method, to give the hydrograph at output. From that first pair of input and output hydrographs we obtained the transfer function using spectral techniques developed above. Now considering the second pair of hydrographs, with different input, we convolved the transfer function with the inflow and compared the resulting outflow with that obtained from the corresponding accurate simulation.

The model river, of infinite width (that does not matter for our purposes), was 100 km long, had a slope of 10^{-4}, with Manning's $n = 0.04$. The initial flow per unit width was $q = 1\,\mathrm{m}^2\,\mathrm{s}^{-1}$, with an initial depth of 2.30 m. In each case a flood was introduced at the upstream end and the governing long wave equations were solved over a six-day period. The traditional point-wise method and three spectral approximation methods were applied: Fourier with $M = 6$, the polynomial series with $M = 10$, and splines with 12 intervals, giving close agreement with each

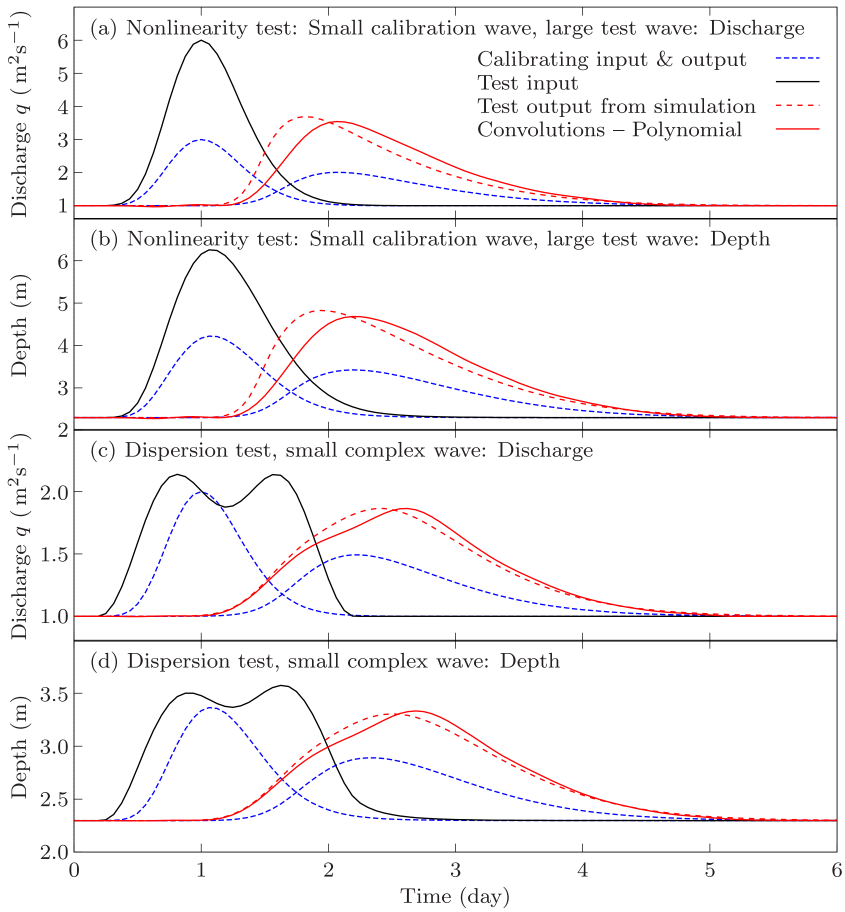

Fig. 5. Results from system identifications for two different combinations of input and output for identification and markedly different input and output for testing.

other. Only the polynomial spectral results are shown in Fig. 5. It is noteworthy that in these problems, the traditional point-wise approach failed badly, giving highly-oscillatory results for the transfer function h_k, similar to those shown in Fig. 3(b)–(d). The reason here may paradoxically have to do with the fact that the hydrograph data show smooth behaviour, when the convolution equations are all similar and hence poorly-conditioned.

The first test was that of the effects of nonlinearity on the accuracy of the linear methods. In the first calibrating case a smoothly-varying flood was introduced with a peak discharge after 1 d of three times the initial flow. The maximum water depth upstream was about 4 m, about twice the initial. Then, for the testing case, a larger flood with a maximum of six times the initial flow was introduced, with a maximum upstream depth of about 6 m, three times the initial.

Both discharge and water level formulations were tested, using data from the same flood event. Results using discharge with a rather larger variation are shown in Fig. 5(a), and show that the magnitude of the downstream flood wave was calculated accurately, but its peak arrived later than that obtained from the accurate simulations, showing that the higher test wave did travel faster than the smaller calibration wave. Part (b) of the figure using surface elevation, plotted as depth show markedly the surprising result that the smaller variation of surface elevation, and hence possible better description using a linear theory has no advantage — the results in parts (a) and (b) are quite similar in the relative deviation of the output from convolution compared with that from accurate numerical simulation.

The second test was more of an examination of the ability of the convolution methods to handle irregularities in the upstream hydrograph (Fenton, 2019, showed that the commonly-used Muskingum method for flood routing is in fact numerically-dispersive and does not handle irregular hydrographs well). Here, a test with a relatively small flood, a calibrating discharge maximum of two times the initial flow was considered. In the testing case, a double-peaked hydrograph with a maximum approximately that of the calibrating case was used. Results are shown in Fig. 5(c) for discharge and (d) for water depth, and it can be seen that the combination of nonlinearity and irregularity has led to finite errors — but again, that results using discharge and depth are very similar.

It seems that linear convolution methods might be able to be used to describe routine river variations, as in Section 5.4, but they are not so accurate for large floods. One could say, however, that the errors shown in these demanding tests are no greater than those of river simulations with unknown geometry and resistance properties, and especially those of the diffusive Muskingum method. They are also probably less than those of rainfall–runoff applications which cannot be tested by accurate simulation.

6. Conclusions

Convolution equations can be used to connect the input and the output of a system such as rainfall and runoff or inflow and outflow of a river reach. In the past, problems of identifying the transfer function, the link between cause and effect, have been widely reported, with large oscillations. Here it is asserted that they are primarily due to the fact that each equation, using a sequence of slowly-changing input numbers, is closely similar to the next, giving a well-known condition for ill-conditioning and hence fluctuations in the solutions. It has been shown that a fundamental solution of such equations is that of wild oscillations of the transfer function, such as has often been found in computations.

Spectral methods have been proposed here to solve the problem, where instead of individual solution values, free to fluctuate, they are expressed as a series of given continuous functions, where the problem is now to determine the coefficients of those functions. Three different forms of series were considered: Fourier, polynomials, and piecewise continuous splines. The equations for each were found to be well-conditioned, and solutions obtained were smooth, bounded, and enabled a certain amount of physical interpretation, previously confused by the erratic nature of results.

The methods have been applied to several problems. For typical rainfall–runoff ones, quite satisfactory results could be obtained. Considering wave and flood propagation problems, good results were obtained for comparison with a large field experimental study involving surges downstream of a power station. Some demanding tests have been made of the ability of the present linear methods to handle problems of nonlinearity in flood routing, where the calibrating/identifying event might have different size and nature from that which it is intended to simulate. Results were not as accurate as desired. However as a first approximation, using methods that require no knowledge of the stream geometry or its nature, the present methods might be useful.

One incidental advantage for river problems is that the methods can be applied to either water level (stage) or discharge hydrographs. In particular, if the former is used, there is no need for the use of rating curves or the effects of unsteadiness on them. Often, water level is the most important quantity anyway.

Practical deductions and implications are:

- Spectral techniques allow automatic treatment of deconvolution problems for arbitrarily varying input and output. The methods are approximate but paradoxically give more useable results.
- One should use only convolution equations from a later point such that while all values of the input sequence are used, only output values and equations are considered such that the convolutions do not have to be truncated by otherwise reaching back beyond the initial point.
- The point values of the transfer function can be approximated by relatively short series of simple continuous functions: Chebyshev polynomials, easily expressed in terms of elementary functions, are recommended. The number of unknowns, and computational costs, are small. If one chooses too high a level of approximation, worse oscillatory results can be obtained, as the spectral method begins to approximate the traditional point-wise convolution form, showing what is called *over-fitting* in statistics.
- In considering application to flood routing problems, where there is a delay between input and output, there are advantages in incorporating that delay into the convolution and not requiring output contributions until some time after input has commenced. However that is not essential.
- No conclusions or recommendations have been made for the length of the transfer function. That seems to be problem-specific, and possibly has to be determined by trial and error.
- Similarly, none have been found or made for the degree of the polynomial series, or length of the Fourier series, or number of spline intervals. Again, regrettably, trial and error is required, although acceptable results can be obtained with small values.

In the past, the use of convolutions and transfer functions has attracted a certain disfavour in practice, in view of the difficulty of obtaining reliable results from deconvolution. If that problem is solved, it is noteworthy that they seem to be a robust and reliable means of approximation. In the course of this work, the author developed a new respect for them.

CRediT authorship contribution statement

John D. Fenton: Conceptualization, Formal analysis, Methodology, Project administration, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The author acknowledges TU Wien Bibliothek for financial support for publication through its Open Access Funding Programme.

References

Abramowitz, M., Stegun, I.A., 1965. Handbook of Mathematical Functions. Dover, New York, https://personal.math.ubc.ca/~cbm/aands/.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology. McGraw-Hill, New York.

Cudennec, C., 2005. Unit hydrograph. In: Lehr, J., Keeley, J., Lehr, J., Kingery, T.B. (Eds.), Water Encyclopedia – Surface and Agricultural Water. Wiley-Interscience, pp. 221–222.

Dooge, J.C.I., 1973. Linear Theory of Hydrologic Systems. Technical Bulletin 1468, United States Department of Agriculture, Washington, D.C, https://fdocuments.net/document/linear-theory-of-hydrologic-systems.html.

Doyle, W.H., Shearman, J.O., Stiltner, G.J., Krug, W.R., 1984. A Digital Model for Streamflow Routing By Convolution Methods. Water-Resources Investigations Report 83-4160, U.S. Geological Survey.

Faye, R.E., Cherry, R.N., 1980. Channel and Dynamic Flow Characteristics of the Chattahoochee River. Buford Dam to Georgia Highway 141, Water Supply Paper 2063, U.S. Geological Survey.

Fenton, J.D., 2018. On the generation of stream rating curves. J. Hydrol. 564, 748–757.

Fenton, J.D., 2019. Flood routing methods. J. Hydrol. 570, 251–264.

Fenton, J.D., 2023. Flood routing using models based on input and output data. In: M.B. Kalinowska, M.M. Mrokowska, Rowiński, P.M. (Eds.), Proc. 40th International School of Hydraulics. Springer, Poland.

Flood Studies Report, 1975. Flood Routing Studies, vol. 3, Natural Environment Research Council, London.

Goring, D.G., 1984. Flood routing by a linear systems analysis technique. J. Hydrol. 69, 59–76.

HEC-HMS, 2023. Technical reference manual of the hydrologic modeling system. https://www.hec.usace.army.mil/confluence/hmsdocs/hmstrm.

Keefer, T.N., 1974. Desktop computer flow routing. J. Hydraul. Div. ASCE 100 (HY7), 1047–1058.

Keefer, T.N., 1976. Comparison of linear systems and finite difference flow-routing techniques. Water Resour. Res. 12 (5), 997–1006.

Keefer, T.N., McQuivey, R.S., 1974. Multiple linearization flow routing model. J. Hydraul. Div. ASCE 100 (HY7), 1031–1046.

Kjeldsen, T.R., 2007. Flood Estimation Handbook – Supplementary Report (1) the Revitalised FSR/FEH Rainfall-Runoff Method. Centre for Ecology and Hydrology, Wallingford.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C, second ed. Cambridge.

Sauer, V.B., 1971. The Unit Hydrograph – a Technique for Routing Reservoir Releases. U.S. Geological Survey Professional Paper 750-B, pp. B259–B264.

Sauer, V.B., 1973. Unit-response method of open-channel flow routing. J. Hydraul. Div. ASCE 99 (HY1), 179–193.

Shaw, E.M., 1994. Hydrology in Practice, third ed. Chapman and Hall.

Sonnenwald, F., Mark, O., Stovin, V., Guymer, I., 2021. Predicting manhole mixing using a compartmental model. J. Hydraul. Eng. 147 (12), 04021046.

Stovin, V.R., Guymer, I., Chappell, M.J., Hattersley, J.G., 2010. The use of deconvolution techniques to identify the fundamental mixing characteristics of urban drainage structures. Water Sci. Technol. 61 (8), 2075–2081.

Yang, Z., Han, D., 2006. Derivation of unit hydrograph using a transfer function approach. Water Resour. Res. 42 (1), W01501.

Zhao, B., Tung, Y.-K., Yang, J.-C., 1995. Estimation of unit hydrograph by ridge least-squares method. J. Irrig. Drain. Eng. 121, 253–259.