19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011
http://mssanz.org.au/modsim2011

Statistical Estimation of Total Discharge Volumes

D.E. Pagendam [a], **A.H. Welsh** [b]

[a] *CSIRO Mathematics, Informatics and Statistics, EcoSciences Precinct, PO Box 2583, Brisbane, QLD 4001*
[b] *Center for Mathematics and Its Applications, Australian National University*
Email: dan.pagendam@csiro.au

Abstract: Calculating the volumes of water discharged by streams is becoming increasingly important in water accounting and deciding how much water to allocate to competing uses. Water accounting is particularly important in Australia, as the driest inhabited continent and also in the face of potential impacts of a changing climate. Stream networks all over the world are littered with gauging stations, which take regular measurements of steam flow in order to help natural resource managers make decisions regarding water allocation. Estimating total discharge volumes is also of utmost importance when estimating pollutant loads from catchments.

In order to calculate the total discharge volume, one must integrate the hydrograph (the graph of stream flow with time) over the period of interest. The simplest method to perform the integration is a trapezoidal scheme, however this fails to account for a number of sources of uncertainty inherent in the hydrograph, namely: (i) what happens between the discrete observations; (ii) gauging stations measure water height and flow is estimated using a rating curve between height and flow; and (iii) there are measurement errors associated with the height data recorded at gauging stations. We present a Monte Carlo method that employs: (i) nonparametric stochastic differential equations (SDEs) to bridge the gaps between discrete observations; and (ii) the Weighted Nadaraya-Watson estimator to estimate the conditional distribution of log-flow given water height. The output of the method is an ensemble of hydrographs that are faithful to the observed data, but incorporating these uncertainties/errors. Integrating the members of this ensemble gives rise to a distribution for the total discharge volumes and properly accounts for the imperfect information available.

We demonstrate the methods using hydrographic data from Obi Obi Creek in the Mary River catchment, Queensland, Australia and examine the uncertainty inherent in total discharges when integrating over a single month and over an entire year. We also introduce an artificial gap of 375 days into the hydrograph and demonstrate how well our simulated diffusions replicate the dynamics of stream flow. Whilst our Monte Carlo method is useful for estimating total discharge volumes, the nonparametric SDEs used also appear to have good potential as stochastic rainfall-runoff models in their own right.

Keywords: hydrograph, stochastic differential equation, diffusion, uncertainty

1 INTRODUCTION

Estimating total discharge of streams is an important part of water resource management. In deciding how best to allocate volumes of water to the likes of irrigators, drinking water supplies and environmental flows, it is imperative to study historical hydrographic records. If one was interested in the total annual flow of a stream, the hydrograph (the record of stream flow over time) could be integrated for each year yielding a distribution of values, from which summary statistics can be computed. At first glance, this is a simple task, however, integrating the hydrograph is in itself a delicate problem and care should be taken to account for a number of sources of uncertainty that are typically inherent in the data. The purpose of this work is to outline a statistical method for integrating hydrographic data in the face of multiple sources of uncertainty, namely: (i) process uncertainty; (ii) rating curve uncertainty; and (iii) measurement uncertainty.

Process uncertainty refers to our lack of knowledge about what transpires between the observations of flow, acknowledging that gauging stations capture only discrete observations of an underlying continuous-time stochastic process. Clearly, such uncertainty decreases as the density of observations increases on the time line and will be of greater importance in situations where discharge is measured daily than where measurements occur with a frequency in the order of minutes. Process uncertainty also becomes increasingly important when there are periods of missing data in a flow series. Such periods are commonplace in hydrographic data and are typically the result of equipment malfunction, vandalism or the lodging of debris in control weirs. Rating curve uncertainty corresponds to errors in the observed flow. Measuring flow directly is expensive and time-consuming and requires careful measurements of channel geometry and stream velocity by experienced personnel. For this reason, a typical gauging station measures only water depth, using a pressure transducer. Hydrographers do record actual stream flow on a regular basis (say monthly) along with water height and use this data to build a (typically nonlinear) relationship between depth and flow, called a rating curve. Using this rating curve, height data collected by gauging stations can be used to estimate flow, but it is imperative to acknowledge the uncertainty surrounding such estimates. Measurement uncertainty can occur in a number of ways, but frequently stems from the measurement precision of instruments. For example, a pressure transducer at a gauging station, measuring water depth with a certain (small) error, might be reported as $\pm$5mm. Each of these three sources of error/uncertainty should be acknowledged and incorporated into our estimates of total discharge when integrating a hydrograph. Doing so presents the practitioner with a distribution of discharge volumes for a period of interest rather than the single number obtained when integrating a hydrograph with, say, a trapezoidal integration scheme.

In order to capture process uncertainty, we model stream flow as a continuous-time stochastic process using stochastic differential equations (SDEs). SDEs have enjoyed widespread use in fields such as finance and econometrics, but surprisingly have had only limited uptake in many other fields, where their uses seem obvious, including hydrology. Some exceptions in the water quality domain include the work of: Finney et al. (1982) who modelled the joint and marginal probability distributions of biological oxygen demand (BOD) and dissolved oxygen (DO) at any point along a river; Curi et al. (1995) who looked at a stochastic model of DO for the Thames River; and Zielinski (1991) who discussed the use of SDEs for modelling BOD. In catchment modelling, Unny (1984); Unny and Karmeshu (1984) examined the use of SDEs as models for the generation of stochastic streamflows. Tsai and Oh (1995) discuss the use of jump-diffusion processes for modelling sediment transport, whilst Marinho et al. (2001) demonstrates how SDEs can be used for modelling volumes discharged into a dam for the purpose of flood control. A review of the subject of SDEs with some additional insight into their applications in water resource engineering is given in Bodo et al. (1987).

Accounting for uncertainty in rating curves has been a focal point of statistical hydrology and a plethora of approaches have been proposed. Petersen-Øverleir (2004) discusses the classical approach of using nonlinear-least squares (NLS) for constructing classical power-law type rating curves, pointing out that this method is only applicable for modelling a very narrow class of variance heterogeneity and that a heteroscedastic maximum likelihood model would be more appropriate. Moyeed and Clarke (2005); Petersen-Øverleir et al. (2009) propose the use of Bayesian methods for fitting rating curves, whilst Ingimarsson et al. (2010) suggest adding a B-spline term to the standard power-law rating curve model within a Bayesian inferential framework. Other approaches proposed include Jansson (1996) who suggests stratifying data into discharge classes for developing sediment rating curves and Ghimire and Reddy (2010) who use a machine learning approach to develop rating curves. In this work we adopt a fully nonparametric approach to fitting rating curves and quantifying uncertainty using kernel estimators, which have seen application in other areas of hydrology (see Moon and Lall (1994); Apipattanavis et al. (2010) for example).

We present novel statistical methodology for estimating total discharge volumes in light of multiple sources of uncertainty. The methods are predominantly nonparametric and data driven, so that minimal analysis is required on behalf of the practitioner. This makes the methods particularly useful for inclusion in software used by water resource managers for water accounting. SDEs are used to model the behaviour of the hydrograph and incorporate inter-observational stochasticity into the integration procedure. The drift and volatility functions of the SDEs are estimated nonparametrically from the observed data, offering a rich class of possible stream behaviour. The error in estimating discharge from a rating curve is quantified using a kernel regression estimator known as the weighted Nadaraya-Watson (WNW) estimator. Rather than use individual rating curves for different periods of time, we incorporate the time of the rating curve data as an independent variable in our model, so that our estimates of discharge are smoothed with respect to both water depth and time.

2 DIFFUSION MODELS FOR STREAM FLOW

One of the most common approaches used to describe a continuous-time, continuous-state stochastic process, X_t, is the SDE (see Klebaner (1998)), which typically has the form $dX_t = \mu(X_t)dt+\sigma(X_t)dB_t$, where $\mu(\cdot)$ is known as the drift function, $\sigma(\cdot)$ is known as the volatility function and B_t is a standard Brownian motion. The solution X_t to such an SDE is known as a *diffusion process*. For many complex systems, including those in hydrology, it is difficult to formulate defensible parametric forms for the drift and volatility functions. There has therefore been a push in recent years towards estimating $\mu(\cdot)$ and $\sigma(\cdot)$ nonparametrically, based on a series of discrete observations from X_t (Aït-Sahalia (1996); Jian and Knight (1997); Bandi and Phillips (2003)).

For modelling stream flow, we adopt the following SDE $dY_t = \mu(Y_t, s_t, r_t)dt + \sigma(Y_t, s_t, r_t)dB_t$, where Y_t is the log-flow process; s_t is a seasonal, deterministic process taking values in $[0, 1]$ and corresponding to the proportion of the year that has transpired; and r_t is a rainfall variable obtained by spatially averaging daily rainfall grids over the catchment area of a gauging station. Bandi and Phillips (2003); Bandi and Moloche (2008) showed that $\mu(\cdot)$ and $\sigma(\cdot)$ can be estimated nonparametrically and these estimators may be both consistent and asymptotically normal (using both infill and long span asymptotics) even if the diffusion process is non-stationary. All that is required to prove consistency and asymptotic normality is that Y_t is *Harris recurrent*, meaning that the process will transition from any subset of the state space to any other subset of the state space infinitely often over an infinite time horizon. This is as benign an assumption as one could hope for and can reasonably be assumed to hold for many hydrological systems of interests. It opens up an extremely rich class of non-stationary, seasonal stochastic processes for modelling the types of complex behaviours observed in stream flow, provided sufficient historical data are available for estimation.

Estimation of the drift and volatility functions uses the Nadaraya-Watson estimators Nadaraya (1964); Watson (1964) based on those in Bandi and Phillips (2003); Bandi and Moloche (2008) and having the form:

$$\hat{\mu}(y, s, r) = \frac{\sum_{i=1}^{n-1} \frac{(Y_{i+1}-Y_i)}{(t_{i+1}-t_i)} K_{h_1}(Y_i - y)K_{h_2}(s_i - s)K_{h_3}(r_i - r)}{\sum_{i=1}^{n-1} K_{h_1}(Y_i - y)K_{h_2}(s_i - s)K_{h_3}(r_i - r)}$$

and

$$\hat{\sigma}^2(y, s, r) = \frac{\sum_{i=1}^{n-1} \frac{(Y_{i+1}-Y_i)^2}{(t_{i+1}-t_i)} K_{h_1}(Y_i - y)K_{h_2}(s_i - s)K_{h_3}(r_i - r)}{\sum_{i=1}^{n-1} K_{h_1}(Y_i - y)K_{h_2}(s_i - s)K_{h_3}(r_i - r)}.$$

Here, $K_h(\cdot)$ denotes a continuous, bounded and symmetric kernel function with bandwidth h. For our purposes, we choose $K_{h_1}(a) = \frac{1}{h^2\sqrt{2\pi}}\exp(-\frac{a^2}{2h^2})$ (the Gaussian kernel) and $K_{h_2}(s_i - s) = \frac{\exp[\cos(s_i-s)/h_2^2]}{h_2\pi I_0(1/h_2^2)}$ (the Von Mises kernel), where I_0 is the modified Bessel function of order 0. We employ the von Mises kernel, whose domain is the circle, to ensure that observations made late in the year have influence on those early in the year and vice versa. Note that we have chosen the same three bandwidths for the estimation of the drift and diffusion functions. This is not entirely necessary, but drastically reduces the dimension of the bandwidth search space when performing cross-validation.

Successful implementation of our diffusion model relies on the appropriate choice of three bandwidths. These bandwidths were selected using 10-fold cross-validation, removing blocks of the hydrograph with each fold. Using the remaining data (i.e. not in the block), drift and diffusion functions were estimated and then, in conjunction with our SDE, used to simulate m hydrographs over the missing block. For an observation of flow y_t at time t, the values of the m diffusions (from the missing block in which y_t resides) were used to generate a kernel density estimate $f_t(y_t|\boldsymbol{h})$. Bandwidths were chosen to maximise $\text{argmax}_{\boldsymbol{h}\in\mathcal{H}} \sum_{i=1}^n \log[f_i(y_i|\boldsymbol{h})]$, the log-likelihood of the observed hydrograph under the diffusion model using a grid search.

3 RATING CURVE UNCERTAINTY

We use the WNW estimator (Hall et al., 1999; Cai, 2001) to estimate, nonparametrically, the distribution of flow given an observation of stream depth. This estimator has a number of desirable properties, namely, it reproduces the superior bias properties of locally linear estimators, like the estimator proposed by Yu and Jones Yu and Jones (1998), but additionally, unlike the Yu-Jones estimator, it always returns a valid distribution function. For our purposes herein, we define the WNW estimator of the distribution function as

$$\hat{F}(Y_t = y|D_t = d, T = t) = \frac{\sum_{i=1}^r p_i(d,t,\mathbf{Y},\mathbf{D},\mathbf{T})K_{h_4}(d-D_i)K_{h_5}(t-T_i)\mathbb{I}(Y_i \le y)}{\sum_{i=1}^r p_i(d,t,\mathbf{Y},\mathbf{D},\mathbf{T})K_{h_4}(d-D_i)K_{h_5}(t-T_i)},$$

where y, d and t are the log-discharge, stream depth and observation time respectively. $\mathbf{Y}, \mathbf{D}$ and $\mathbf{T}$ are r-dimensional vectors containing the observations of log-discharge, stream depth and the observation times respectively, that are used to build the rating curves for the gauging station. The functions $K_{h_4}(\cdot)$ and $K_{h_5}(\cdot)$ are Gaussian kernel functions (as defined above) with bandwidths h_4 and h_5 respectively and $\mathbb{I}(\cdot)$ is the indicator function, which takes the value of one if the expression in brackets is true and zero otherwise. Finally, the $p_i(d,t,\mathbf{Y},\mathbf{D},\mathbf{T})$ are a set of additional weights, that are a function of the rating curve data as well as the time (t) and depth (d) at which the distribution is being computed. These weights have the property that each $p_i \ge 0$, $\sum_{i=1}^n p_i = 1$, in addition to

$$\sum_{i=1}^r (D_i - d)p_i(d,t,\mathbf{D},\mathbf{T})K_{h_4}(d-D_i)K_{h_5}(t-T_i) = 0 \quad \text{and}$$
$$\sum_{i=1}^r (T_i - t)p_i(d,t,\mathbf{D},\mathbf{T})K_{h_4}(d-D_i)K_{h_5}(t-T_i) = 0.$$

The weights satisfying these constraints are equal to $p_i = M_i/\sum_{i=1}^r M_i$, where

$$M_i = \{1 + \lambda_1 K_{h_4}(d-D_i)K_{h_5}(t-T_i)(d-D_i) + \lambda_2 K_{h_4}(d-D_i)K_{h_5}(t-T_i)(t-T_i)\}^{-1}.$$

and λ_1 and λ_2 are Lagrange multipliers. The $\{p_i\}$ satisfying the above constraints are not uniquely defined and the accepted approach is to choose λ_1 and λ_2 in order to maximise $\sum_{i=1}^r \log(p_i)$.

The procedure outlined above allows us to estimate the distribution function of stream flow conditional on the observed stream depth at any time in the gauging station's history. The incorporation of time into the estimator produces a conditional distribution function that varies smoothly through time to incorporate observed changes in the depth-flow relationship.

The bandwidths, h_4 and h_5 were chosen by fitting a kernel regression between log-flow, depth and time using the function 'npregbw' within the 'np' library of the R statistical package (R Development Core Team, 2011).

4 MONTE CARLO SIMULATIONS

In order to assess the distribution of the total discharge volume over a period of interest, we wish to simulate N hydrographs that are faithful to the observed data, but incorporating: (i) process uncertainty; (ii) rating curve uncertainty; and (iii) measurement error. This Monte Carlo scheme is described by the following steps.

Sampling from the Distribution Function of Total Discharge Volumes

1. Set $N = 1000$ or some other appropriately large value, defining the number of simulations to perform and set $i = 1$.
2. Generate n uniformly distributed random variables, $(U_{i,1}, \ldots, U_{i,n})$ over the interval $[0, 1]$ and n uniformly distributed random variables, $(W_{i,1}, \ldots, W_{i,n})$ over the interval $[-0.005, 0.005]$.
3. For each observation $d_j \in \mathbf{d}$ $(1 \leq j \leq n)$, generate a sample of stream depth (in metres) as $D_{i,j} = d_j + W_{i,j}$.
4. For each element $(t_j; D_{i,j}) \in (\mathbf{t}; \mathbf{D}_i)$, calculate $\hat{F}(\cdot|D = D_{i,j}, T = t_j)$ and obtain samples of log-flow $Y_{i,j}$ by inverting the estimated conditional distribution function: $Y_{i,j} = \hat{F}^{-1}(U_{i,j}|D = D_{i,j}, T = t_j)$.
5. For each pair of successive log-flow samples $(Y_{i,j}, Y_{i,j+1})$, generate a diffusion bridge $\breve{\mathbf{Y}}_i^{j,(j+1)}$, having B_j time intervals of length Δt, beginning at $Y_{i,j}$ and ending at $Y_{i,j+1}$.
6. Compute the total discharge volume as the integral of the simulated diffusion process using
$$V_i = \frac{1}{2}\sum_{j=1}^{n-1}\sum_{k=1}^{B_j}\frac{(t_{j+1} - t_j)}{B_j}\left[\exp(\breve{Y}_{i,k}^{j,(j+1)}) + \exp(\breve{Y}_{i,(k+1)}^{j,(j+1)})\right]$$
7. If $i < N$, then $i = i + 1$ and go to step 4, otherwise go to step 8.
8. Return $\mathbf{V} = (V_1, \ldots, V_N)$ as an independent and identically distributed sample from the distribution function of the total discharge volume.

The diffusion bridges referred to in step 5 are realisations of the diffusion process that start and finish at particular values of log-flow. There are achieved here by a simple rejection sampling method, whereby a trajectory is retained as a diffusion bridge if it begins at a log-flow of, say, a and ends close to the endpoint b, say, within $[b - \epsilon, b + \epsilon]$, where $\epsilon > 0$ is small. The diffusions were simulated according to an Euler-Maruyama scheme $Y_{i+1} = Y_i + \mu(Y_i, r_i, s_i)\Delta t + \sigma(Y_i, r_i, s_i)\sqrt{\Delta t}Z_i$, where Z_i is a standard normal random variate.

We apply our Monte Carlo method to hydrographic data from Gardener's Falls at Obi Obi Creek in the Mary River catchment, Queensland, Australia. We estimated the drift and diffusion functions using daily flow data over the period 1/1/2000 - 11/4/2010 and estimated the total discharge volumes for the month of January 2009 and the entire year 2009. The probability distributions obtained were then compared with the discharge volumes obtained under trapezoidal integration of the daily flow data when accounting for no uncertainties.

As an additional illustration of the usefulness of these methods, we imposed an artificial gap of 375 days in the hydrographic data, similar to those that occur when a station breaks down, in order to assess the ability of our method to statistically infill the hydrograph.

Table 1. Total Discharge Volumes based on trapezoidal integration and our Monte Carlo method.

Period	V_{trap}	$\text{Mean}(\mathbf{V})$	$\text{Min}(\mathbf{V})$	$\text{Max}(\mathbf{V})$	$Q_{0.05}(\mathbf{V})$	$Q_{0.95}(\mathbf{V})$
1 Jan 2009 - 31 Jan 2009	0.0259	0.0247	0.0213	0.0288	0.0230	0.0266
1 Jan 2009 - 31 Dec 2009	0.9343	0.9237	0.8611	1.0046	0.8775	0.9676

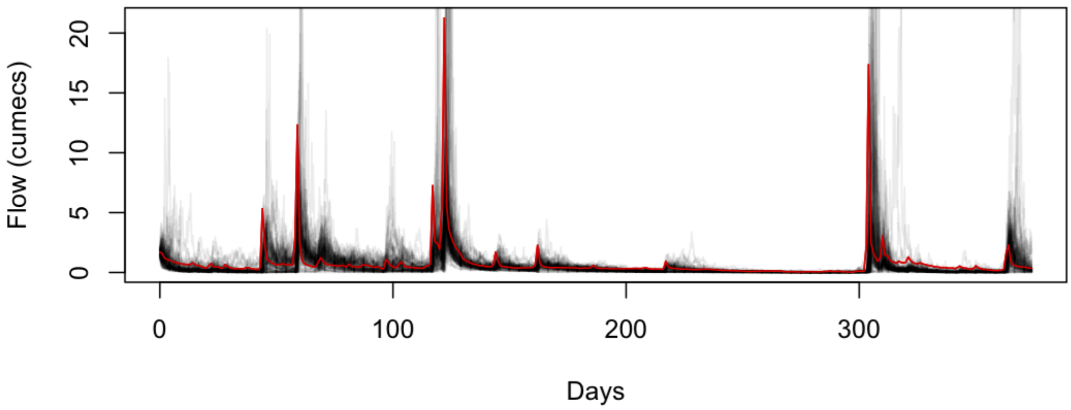

Figure 1. A 375 day gap imposed on the hydrograph (red) and density plot of 100 diffusion bridges used to fill the gap (grey).

5 RESULTS AND CONCLUSIONS

Table 1 shows the total discharge volumes computed using a simple trapezoidal integration (V_{trap}) and the mean, minimum, maximum, 5th percentile ($Q_{0.05}$) and 95th percentile ($Q_{0.95}$) of discharge volumes from 1000 simulated hydrographs (the elements of $\boldsymbol{V}$). There is good agreement between the mean of $\boldsymbol{V}$ and V_{trap} and V_{trap} is contained within the bounds of $[Q_{0.05}(\boldsymbol{V}), Q_{0.95}(\boldsymbol{V})]$ for both time periods examined. We note that the amount of variability in the distributions of flow is not insignificant when examined for a single month and for a year.

Figure 1 also demonstrates that our nonparametric diffusion bridges do a respectable job of infilling even very long gaps in the hydrograph. The SDEs appear to emulate the behaviour of the stream well, capturing the quickflow events, the recession curve and baseflow conditions, based solely on rainfall and seasonal data. The results suggest that these SDEs also have potential as a class of 3-parameter stochastic rainfall-runoff models in other application areas.

ACKNOWLEDGEMENT

This research was supported by CSIRO's Water for a Healthy Country Flagship and a CSIRO OCE Postdoctoral Research Fellowship. We thank the Queensland Department of Environment and Resource Management (DERM) for supplying hydrographic and rainfall data.

REFERENCES

Aït-Sahalia, Y. (1996). Nonparametric pricing of interest rate derivative securities. *Econometrica 64*, 527–560.

Apipattanavis, S., B. Rajagopalan, and U. Lall (2010). Local polynomial-based flood frequency estimator for mixed population. *Journal of Hydrologic Engineering 15*(9), 680–691.

Bandi, F. and G. Moloche (2008). On the functional estimation of multivariate diffusion processes. http://faculty.chicagobooth.edu/federico.bandi/research/Bandi_Moloche.pdf. [Online; accessed 1-February-2011].

Bandi, F. and P. Phillips (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica 71*(1), 241–283.

Bodo, B., M. Thompson, and T. Unny (1987). A review on stochastic differential equations for applications in hydrology. *Stochastic Hydrology and Hydraulics 1*, 81–100.

Cai, Z. (2001). Weighted Nadaraya-Watson regression estimation. *Statsitics and Probability Letters 51*, 307–318.

Curi, W., T. Unny, and J. Kay (1995). A stochastic physical system approach to modeling river water quality. *Stochastic Hydrology and Hydraulics 9*, 117–132.

Finney, B., D. Bowles, and M. Windham (1982). Random differential equations in river water quality modelling. *Water Resources Research 18*(1), 122–134.

Ghimire, B. and M. Reddy (2010). Development of stage-discharge rating curve in river using genetic algorithms and model tree. *International Workshop on Advances in Statistical Hydrology, May 23-25 Taormina, Italy*.

Hall, P., R. Wolff, and Q. Yao (1999). Methods for estimating a conditional distribution function. *Journal of the American Statistical Association 94*(445), 154–163.

Ingimarsson, K., B. Hrafnkelsson, S. Gardarsson, and A. Snorrason (2010). Bayesian discharge rating curves based on b-spline smoothing functions. *Hydrology and Earth System Sciences Discussions 7*, 2747–2780.

Jansson, M. (1996). Estimating a sediment rating curve of the Reventazon River at Palomo using logged mean loads within discharge classes. *Journal of Hydrology 183*, 227–241.

Jian, J. and J. Knight (1997). A nonparametric approach to the estimation of diffusion processes. *Econometric Theory 13*(5), 615–645.

Klebaner, F. (1998). *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London.

Marinho, G., M. Fragoso, and A. Carneiro (2001). A stochastic approach to the flood control problem. *Applied Mathematical Modelling 25*, 499–511.

Moon, Y.-I. and U. Lall (1994). Kernel quantile function estimator for flood frequency analysis. *Water Resources Research 30*(11), 3095–3103.

Moyeed, R. and R. Clarke (2005). The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources 28*, 807–818.

Nadaraya, E. (1964). On estimating regression. *Theory of Probability and Its Applications 10*, 186–190.

Petersen-Øverleir, A. (2004). Accounting for heteroscedasticity in rating curve estimates. *Journal of Hydrology 292*, 173–181.

Petersen-Øverleir, A., A. Soot, and T. Reitan (2009). Bayesian rating curve inference as a streamflow data quality assessment tool. *Water Resources Management 23*, 1835–1842.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Tsai, C. and J. Oh (1995). A stochastic diffusion jump model of suspended sediment transport in hydrologic extreme events. *Stochastic Hydrology and Hydraulics 9*, 117–132.

Unny, T. (1984). Numerical integration of stochastic differential equations in catchment modeling. *Stochastic Hydrology and Hydraulics 20*(3), 360–368.

Unny, T. and Karmeshu (1984). Stochastic nature of outputs from conceptual reservoir model cascades. *Journal of Hydrology 68*, 161–180.

Watson, G. (1964). Smooth regression analysis. *Sankhya Series A 26*, 101–116.

Yu, K. and M. Jones (1998). Local linear quantile regression. *Journal of the American Statistical Association 93*, 228–237.

Zielinski, P. (1991). On the meaning of randomness in stochastic environmental models. *Water Resources Research 27*(7), 1607–1611.