A Model for the Frequency of Extreme River Levels Based on River Dynamics

Stephen Chick* John Shortle† Pieter van Gelder‡
Max B. Mendel§

August 9, 1996

Abstract

A new model for predicting the frequency of extreme river levels is proposed which encapsulates physical knowledge about river dynamics. The central idea is the use of continuous time stochastic processes that use hydrological equations and ergodic theory to model extreme events, rather than relying on statistical fits of classical models to local maximum data. A simple example shows how changes in discharge characteristics changes the extreme river level frequencies. Solutions are provided for special cases, and directions for more general techniques are provided.

Keywords: Extreme river flooding, flood statistics, flood frequency, ergodic theory, engineering probability, Markov-modulated fluid source, Chézy's equation.

1 Introduction

Two major purposes for flood prediction are (1) short-term evacuation purposes and (2) long-term flood frequency analysis for flood defense system design. At present, the analysis for each purpose is quite different. Short-term prediction uses extensive physical knowledge of continuous-time river catchment properties, including river dynamics, prior and predicted weather over the catchment, and run-off models. The FLOFOM model described by Berger [1] and used by

*IOE Dept., University of Michigan, 130 IOE Building, 1205 Beal Avenue, Ann Arbor, Mich 48109

†UC Berkeley, Department of Industrial Engineering and Operations Research.

‡Technical University of Delft, Holland, Department of Civil Engineering.

§UC Berkeley, Department of Industrial Engineering and Operations Research. Research partially supported by the Army Research Office (DAAL03-91-G-0046) grant to the University of California at Berkeley.

the Rijkswaterstaat (Dutch Ministry of Transport and Water Defense) to analyze the river Maas (Meuse) is an example of such a model. Long-term flood prediction, on the other hand, has traditionally fit peak flood levels to statistical models such as the log-Pearson III [2] or Gumbel [3]. For example, Fig. 1 shows statistical fits to flood data from the Rhine [4]. Cunnane [5] provides a survey of these and other statistical models and indicates that goodness of fit tests (χ^2, Kolmogorov-Smirnov, moment-ratio diagrams) for selecting a particular model are inconclusive. Difficulty in comparing hydrological parameters (run-off coefficients, river discharge equations) with statistical parameters (slope, intercept) is one contributor to this inconclusiveness.

In this paper, a new model for predicting the frequency of extreme river levels is proposed which encapsulates physical knowledge about river dynamics, including formulas which describe river discharge. For simplicity, assume the equation for river discharge is given by the widely accepted Chézy equation, a power law, although the theory permits more general formulations. The resulting model therefore makes an explicit connection between hydrological and statistical parameters. The model accounts for the river dynamics at a given location by modeling both how water gets into the river (via upstream tributaries) and how water leaves (discharge modeled by Chézy's equation). Although the simplified physical model presented in Sec. 2 makes several rough approximations (using memoryless properties and Chézy's equation for approximating discharge), some insights may still be gained. The mathematics needed for calculating long-term flood frequencies is presented in Sec. 3. Calculations for several test cases are given in Sec. 4, where it is shown how change in the discharge directly changes the water level for floods with the same return period.

We show that the discharge rates directly affect the shape of the curve which relates flood frequency and flood volume. This is illustrated by the third test case, where the power in the power law of Chézy's equation is varied. In particular, the log-linear relationship commonly assumed is not obtained.

Solution techniques for the general case are presented in Sec. 5.

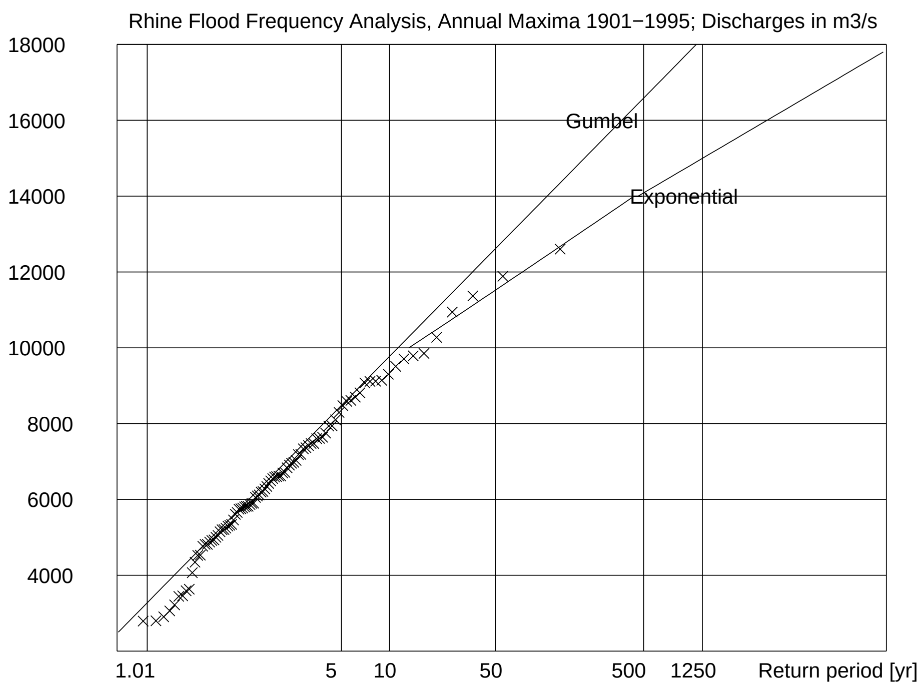

Figure 1: Empirical river discharge frequencies of the Rhine and fitted distributions.

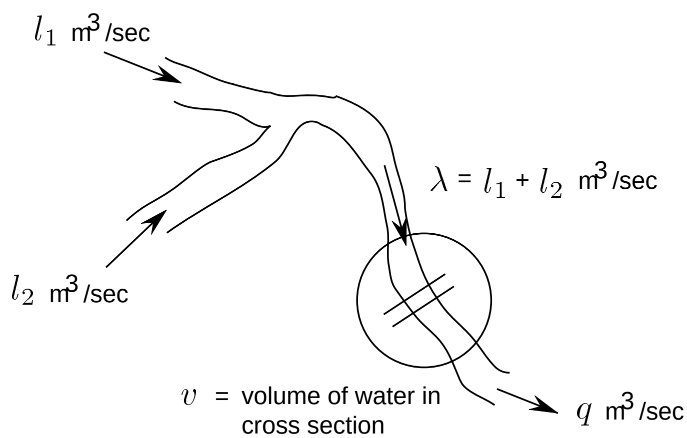

Figure 2: A simplified physical model for the river. At the given cross-section of unit width, the river level is determined by input from $K = 2$ major tributaries and the rate of discharge.

2 Physical model for river system

Short-term flood prediction makes extensive use of physical modeling of the river system. We seek to use a physical model for the river system as part of the long-term prediction model. The main components of a physical model of a river at a particular unit-width cross-section are descriptions of the discharge rate, and the input rate of water from upstream sources (Fig. 2). Because these models may be complex, we propose the following simplified model for how water enters and leaves a river in order to better understand the viability of a physical-model based approach to long-term flood prediction. The physical model presented here is used to calculate the frequency of floods in later sections.

The rate of discharge, $q(v)$, from a unit-width cross-section of a river is a function of the geometry of the cross-section, and the volume of water v in that cross-section. Thus, the discharge has a functional form which is different for the many geometries encountered in practice, such as those seen in Fig. 3.

A standard, accepted description of the rate of discharge is given by Chézy's equation,

$$q(v) = av^b \tag{1}$$

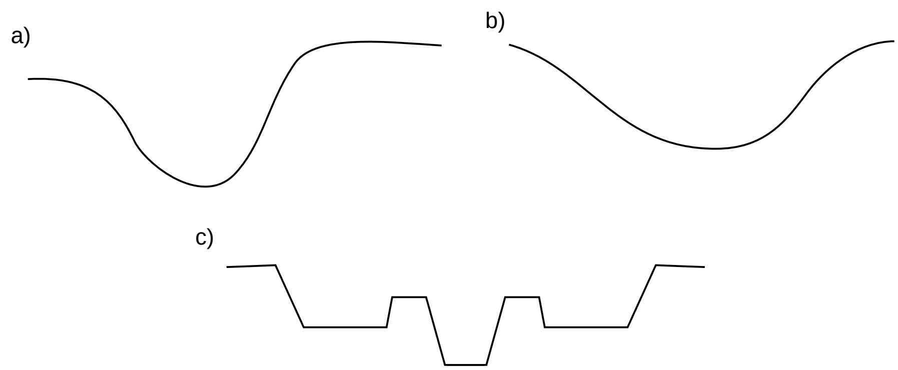

Figure 3: Different geometries for river cross-sections.

for some constants a and b, and q has units m^3/sec.[1] We will use Chézy's equation to describe discharge in this paper, although all the results are valid for more general functions for discharge as well.

The other major feature of a physical model, the rate of water input, is assumed to be determined by the cumulative effect of upstream tributaries. Suppose that there are K major tributaries whose catchments affect the cross-section under consideration, and that the flow time from tributary i to the river cross-section in question is approximately δ_i. Then

$$\lambda(t) \approx \sum_{i=1}^{K} l_i(t - \delta_i) \tag{2}$$

where $\lambda(t)$ is the rate of water input, $l_i(t)$ is the discharge in m^3/sec from tributary i at time t (Fig. 2). Eq. 2 assumes additivity. This provides a first-order approximation for river flows. More sophisticated models which account for non-linearity and 'storage effects' are possible future refinements.

Short term prediction models generally have sophisticated models for relating the weather, runoff, and other factors to l_i. The number of parameters used by short-term models is quite large. For purposes of illustrating ideas,

[1] Many developments have q as a function of the river height h. The present formulation, which approximates volume as a power law in height, is more convenient for the development below.

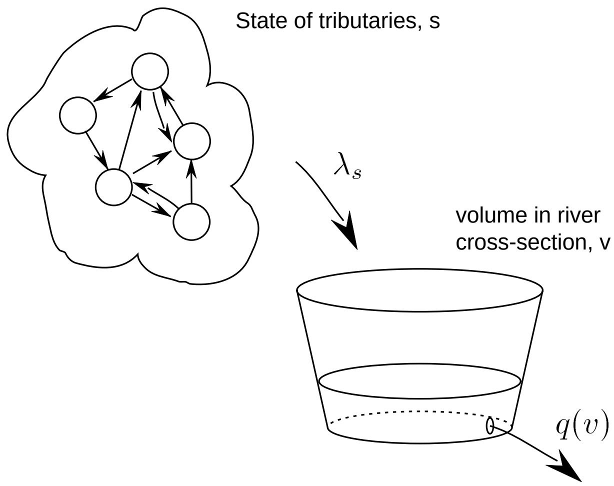

Figure 4: Conceptual model for river: The state of the tributaries s is modeled by a continuous time Markov chain (CTMC). The CTMC state determines the rate of input λ_s from tributaries. Chézy's equation for discharge $q(v)$.

a simplified model is presented. We model the state of the tributaries as a Markov-modulated fluid source (MMFS). The MMFS is commonly used to gain insights into information flow in high-speed communications networks [6].

The basic components of the MMFS are the state and fluid rate. In the present work, the state s is defined to be the array of discharges from the K tributaries. The fluid rate is defined to be the sum of the discharges from the tributaries, and is the input rate of water into the river cross-section.

We make a necessary generalization for the MMFS in order to apply it to the flood-frequency problem. In communications, the MMFS generally assumes a constant bit rate, with the choice $q(v)$ = constant bit rate. The present work makes an extension by permitting $q(v)$ to be a non-constant function, such as Chézy's equation.

The use of the MMFS to model the river system requires two simplifications of the physical model. First, the MMFS that the l_i form a discrete set, $l_i \in \mathcal{L}$. For a particular river system, one might set $\mathcal{L} = \{0, 25, 50, \ldots, \bar{q}\}$ (increments not necessarily constant). The maximal discharge $\bar{q}$ is chosen to be greater than the largest imaginable discharge from the tributaries (this insures the number of states is finite). With this assumption, the effect of river catchment at time t on the river cross-section is described by $S(t) = (l_1(t-\delta_1), \ldots, l_K(t-\delta_K)) \in \mathcal{L}^K$. We call $S(t)$ the state of the tributaries at time t.

Set $N = |\mathcal{L}|^K$ to be the number of possible states of the tributaries. Number the states from $s = 1, \ldots, N$. For each state $s \in \mathcal{L}^K$, the rate of water input λ_s into the river cross-section is

$$\lambda_s = \sum_{i=1}^{K} l_i$$

Set $\Lambda = \mathrm{diag}(\lambda_s)$ to be the diagonal matrix of discretized river inputs.

A second assumption of the MMFS is that the random state transitions are memoryless. Equivalently, the catchment is modeled as a continuous-time Markov chain (see, eg [7]), with state transition matrix

$$M = [\mu_{ij}] \tag{3}$$

with μ_{ij} equal to the state transitions rate from state i to j (when $i \neq j$), and μ_{ii} the negative of the rate out of state i.

$$\mu_{ij} = \begin{cases} \lim_{h\to 0} \mathrm{Prob}\left(S(t+h) = j \mid S(t) = i\right)/h & i \neq j \\ -\sum_{k\neq i} \mu_{ik} & i = j \end{cases} \tag{4}$$

where $i, j \in \mathcal{L}^K$ and $S(t)$ is the state of the tributaries at time t. Thus, the rows add to 0. Under mild technical conditions, the long run probabilities π_i of being in state i satisfy

$$\vec{\pi} M = 0$$

Although M is a large matrix of size $N \times N$, it is sparse. This is due to the fact that rivers do not skip discharge levels instantaneously. For example, with discrete values $25, 50, 75$ m^3/sec, a river cannot skip 50 when changing from 25 to 75 m^3/sec. There are at most K ways to have an increased flow (an increase in one of the K tributaries input flows) and at most K ways to decrease. Thus, each row will have a maximum of $2K+1$ non-zero entries, including the diagonal entry.

This model provides sufficient mathematical structure to calculate the frequency of extreme floods, as detailed further in Sec. 3

3 Mathematical model for flood frequency

This section presents the mathematics needed to calculate the expected return time for a flood based on the model of Sec. 2.

The recurrence time $T(v)$ for a flood with volume v in the cross-section, is given by

$$T(v) = \frac{D(v)}{P(v)} \tag{5}$$

where $D(v)$ is the expected duration of a flood exceeding $V = v$, and $P(v)$ is the probability of exceeding v at an arbitrary time in steady state. This formula can be derived using results from ergodic theory, as described in the appendix.

We now focus on the calculation of $D(v)$ and $P(v)$.

From the simplifying assumptions above, the state of the river is given by the volume v of water in the cross-section of interest, and the state of the tributaries $s \in \mathcal{L}^K$. A differential equation describing the change of probabilities for system states can be given in matrix form as

$$\begin{aligned} \frac{\partial \vec{F}(v;t)}{\partial t} &= \text{rate from water flow} + \text{rate from Markov chain state changes} \\ &= -\left(\Lambda - q(v) I\right) \frac{\partial \vec{F}}{\partial v} + M^T \vec{F} \end{aligned} \tag{6}$$

where $\vec{F}(v;t) = \left(F_1(v;t), \ldots, F_N(v;t)\right)^T$ is a column vector of functions

$$F_s(v;t) = \text{Prob}\{V \leq v,\ \text{tributary state } S(t) = s;\ \text{at time t}\} \tag{7}$$

Eq. 6 has a unique solution whenever the Markov chain M is irreducible and the discharge $q(v)$ is sufficient to prevent permanent flooding [8].

The probability of the river exceeding a flood level v at a 'random time' is then given by

$$P(v) = 1 - \sum_{s=1}^{N} F_s(v),$$

where $F_s(v)$ is the stationary distribution satisfying $\partial \vec{F} / \partial t = 0$. (When no explicit reference to t is given, the stationary solution is implied.) From Eq. 6, $F_s(v)$ can also be determined from

$$\left(\Lambda - q(v) I\right) \frac{\partial \vec{F}}{\partial v} = M^T \vec{F}, \tag{8}$$

a linear, homogeneous, first-order ordinary differential equation.

$D(v)$ is calculated by simulating floods using Eq. 6. Numerical methods for calculating $D(v)$, $P(v)$, and $T(v)$ are discussed in Sec. 5.

4 Sample calculation

The above formulation is sufficient to determine how changes in the river system or discharge rate change the return time of an extreme flood. In this section, three sample calculations for simple systems are given to illustrate how the return periods of extreme floods when the discharge function is changed.

Suppose that either it is raining at λ_r m^3/sec over the catchment, or that no rain enters the river. The basal river flow from springs is λ_2. The river input during rain $\lambda_1 = \lambda_r + \lambda_2$ is the sum of the basal flow and rainfall. Rain lasts an exponential amount of time with mean $1/\mu_{12}$, and periods of no rain are exponential with mean $1/\mu_{21}$. (The exponential model is required for the MMFS.) Then

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} -\mu_{12} & \mu_{12} \\ \mu_{21} & -\mu_{21} \end{pmatrix}$$

The stationary probabilities of rain vs. no rain are $\vec{\pi}_1 = \mu_{21}/(\mu_{12}+\mu_{21})$ and $\vec{\pi}_2 = \mu_{12}/(\mu_{12}+\mu_{21})$, respectively. The average input to the river is $\bar{\lambda} = \frac{\mu_{21}\lambda_1+\mu_{12}\lambda_2}{\mu_{12}+\mu_{21}}$.

The river levels for the cases of two differing powers in Chézy's equation ($b=0$ and $b=1$) are presented. In both cases,

$$\text{Prob}(V \leq v) = \frac{\mu_{21}}{\mu_{12}+\mu_{21}} F_1(v) + \frac{\mu_{12}}{\mu_{12}+\mu_{21}} F_2(v)$$

where $\frac{\mu_{21}}{\mu_{12}+\mu_{21}}$ is the probability of rain at a random time, $F_1(v) = \text{Prob}(V \leq v, \text{ during rain})$, and $F_2(v) = \text{Prob}(V \leq v, \text{ no rain})$.

Insights follow the calculations.

Case 1: Discharge proportional to volume, $q(v) = cv$. The appendix shows that the following functions satisfy the differential equation for the probability of not exceeding a given river level at a random time.

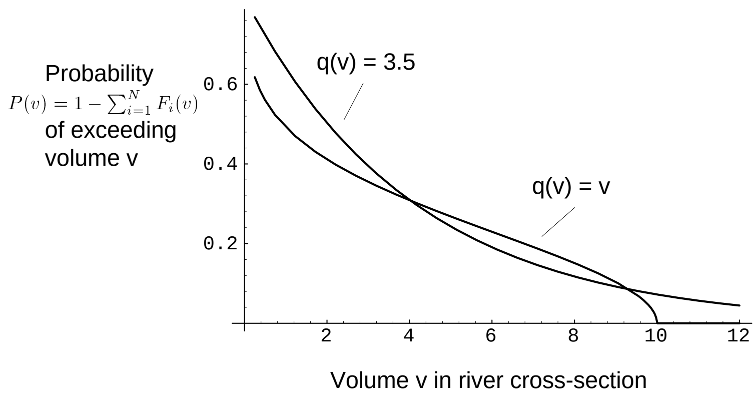

Figure 5: Probability of exceeding a given volume v of water in river cross-section, when $N = 2$, $\lambda_1 = 10$, $\lambda_2 = 0$, $1/\mu_{12} = 2$ days, $1/\mu_{21} = 5$ days.

$$
\begin{aligned}
F_1(v) &= \frac{\mu_{21}}{\mu_{12}+\mu_{21}} I\left(\frac{cv-\lambda_2}{\lambda_1-\lambda_2}, \frac{\mu_{21}}{c}+1, \frac{\mu_{12}}{c}\right) \\
F_2(v) &= \frac{\mu_{12}}{\mu_{12}+\mu_{21}} I\left(\frac{cv-\lambda_2}{\lambda_1-\lambda_2}, \frac{\mu_{21}}{c}, \frac{\mu_{12}}{c}+1\right)
\end{aligned}
\tag{9}
$$

with $v \in (\lambda_2/c, \lambda_1/c)$ and $I(z,a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_{t=0}^{z} t^{a-1}(1-t)^{b-1} dt$ is the incomplete regularized beta function [9].

The expected duration $D(v)$ of a flood of level v or higher with discharge $q(v) = cv$ is

$$
D(v) = \frac{\psi(\lambda_1/c) - \psi(v)}{(\lambda_1 - cv)^{\frac{\mu_{12}}{c}} (cv - \lambda_2)^{\frac{\mu_{21}}{c}}}
\tag{10}
$$

where

$$
\psi(v) = \frac{(\lambda_1-\lambda_2)^{(\mu_{12}+\mu_{21})/c}}{c} \times \left[B\left(\frac{cv-\lambda_2}{\lambda_1-\lambda_2}, \frac{\mu_{21}}{c}+1, \frac{\mu_{12}}{c}\right) + B\left(\frac{cv-\lambda_2}{\lambda_1-\lambda_2}, \frac{\mu_{21}}{c}, \frac{\mu_{12}}{c}+1\right)\right],
$$

where $B(z,a,b) = \int_{t=0}^{z} t^{a-1}(1-t)^{b-1} dt$ is the incomplete beta function.

Case 2: Discharge constant, $q(v) = d$ when $v > 0$. Although no river actually discharges c m^3/sec when there is water and does not discharge when there is none, this special case illustrates how the river levels are affected by

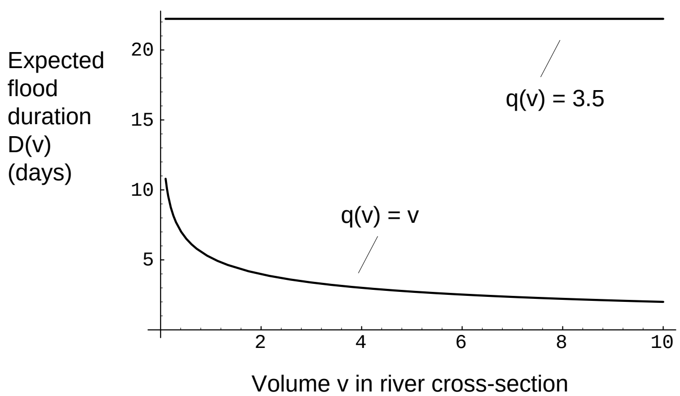

Figure 6: Expected flood duration $D(v)$ same parameters as in Fig. 5.

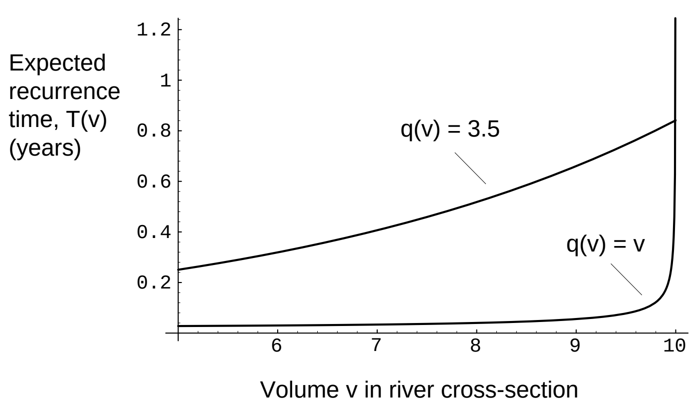

Figure 7: Expected flood recurrence time $T(v)$ (Eq. 5) using same parameters as in Fig. 5.

assuming a special form of Chézy's equation. The assumption $d > \bar{\lambda}$ is necessary so that the river can drain itself.

$$\begin{aligned} F_1(v) &= \frac{\mu_{21}}{\mu_{12}+\mu_{21}} - \frac{\mu_{21}}{\mu_{12}+\mu_{21}} e^{\zeta v} \\ F_2(v) &= \frac{\mu_{12}}{\mu_{12}+\mu_{21}} - \frac{\lambda_1 - d}{d-\lambda_2} \frac{\mu_{21}}{\mu_{12}+\mu_{21}} e^{\zeta v} \end{aligned} \tag{11}$$

where $v \geq 0$ and $\zeta = \frac{d(\mu_{12}+\mu_{21})-\mu_{21}\lambda_1-\mu_{12}\lambda_2}{(d-\lambda_1)(d-\lambda_2)}$.

The expected duration $D(v)$ of a flood of level v or higher is constant

$$D(v) = \frac{\bar{\lambda}-\lambda_2}{\mu_{21}(d-\bar{\lambda})} \tag{12}$$

whenever $v > 0$. When $v = 0$, the expected flood time is infinity, since the water level is always 0 or greater. The appendix provides proof.

Comments The probability of observing a flood at a random time, the expected flood duration, and the expected flood recurrance time for these two examples are illustrated in Fig. 5, Fig. 6, and Fig. 7, respectively.

An important point is revealed by considering Fig. 7, which relates water levels with the corresponding flood recurrence time. The curves are shaped quite differently because of the change discharge characteristics. If a traditional statistical model has the same curve shape relating flood level and flood frequency, it may accurately reflect the discharge characteristics of the river. On the other hand, we have not yet determined which values of the Chézy parameters, if any, reflect the shapes of traditional statistical models.

Another observation is that increasing the power b from Chézy's equation ($q(v) = av^b$) drastically reduces the recurrence time for a given flood level. The exception is near the asymptote near flood volume $v = 10$, which is near the bounds where the model is expected to blow up. One way to avoid this blow up is have a much larger number of states so that the recurrence time of interest (say, 1 250 years) would not be near the asymptote.

Case 3: Effect of Chézy power parameter b on the flood probability $P(v)$. Suppose the discharge is given by Chézy's equation, $q(v) = av^b$. Eq. 5 states that flood return time is proportional to flood duration $D(v)$ and $1/P(v)$, where $P(v)$ is the probability of exceeding v at a random time in steady state.

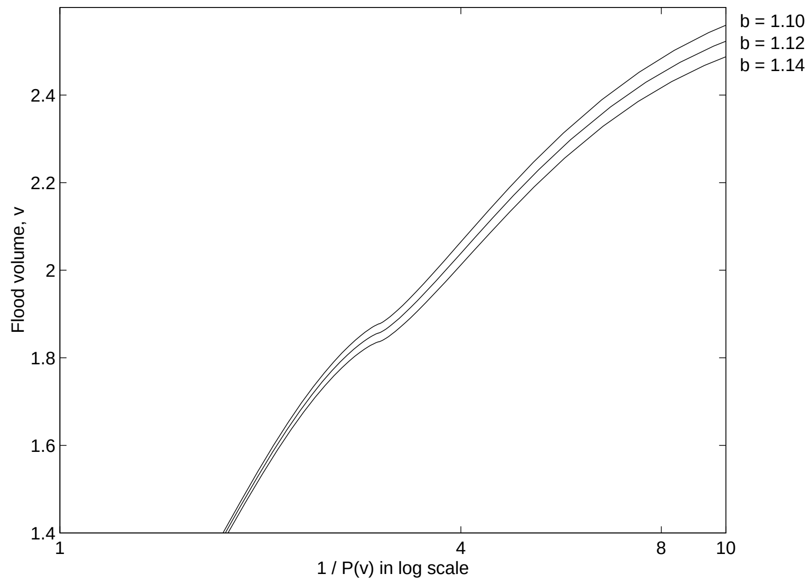

Figure 8: Effect of changes of power parameter b in Chézy's equation on the curve relating flood volume and the steady-state probability of flooding.

Solving Eq. 8 to obtain $1/P(v)$ is numerically simpler than solving Eq. 6 to obtain $D(v)$. In practice, an order of magnitude change of $P(v)$ over a given range of v will not be accompanied by an order of magnitude change in flood duration $D(v)$. Note that small changes in $P(v)$ affect the return time $T(v)$ more than small changes in $D(v)$. We therefore look at how changes in the Chézy parameter b affect $1/P(v)$, since this is likely to be the principal factor influencing $T(v)$.

In particular, assume there are $N = 3$ states, and

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} -0.4 & 0.4 & 0 \\ 0.6 & -1 & 0.4 \\ 0 & 0.6 & -0.6 \end{pmatrix}$$

The rate of increasing to a higher level of river input has been set to a constant 0.4. Chézy's equation is assumed to be $q(v) = 1 \cdot v^b$, for $b = 1.1, 1.12, 1.14$.

Fig. 8 shows how the curve relating v and $1/P(v)$ changes with the Chézy parameter b. Only the portion of the graph away from the maximum and minimum values of v are shown to avoid the asymptotes which are a side-effect of having a finite Markov chain. The axis and scale of the graph have been chosen so that a comparison with the discharge graph of Fig. 1 can be made (assuming $D(v)$ remains relatively constant over the graphed range). The small kink in the graph near $v = 1.85$ is a side-effect of the discretization of the river input rates, λ_i, and may be ignored.

Two observations can be made from this graph. First, better drainage (larger b) results in a reduced probability $P(v)$ of exceeding v, and therefore an increase in $1/P(v)$. Second, and more interestingly, the curves are *not parallel.* The same increase in b causes a comparatively larger change in $P(v)$ for large values of v.

The implication for this example is that increasing b tends to bend the curve downward. A straight line relating flood volume and the $\log P(v)$, as is commonly assumed, is not valid here as b changes.

5 General solution techniques

We turn to the problem of calculating the return periods of extreme floods $T(v)$, where v is the volume of water in the river cross-section of interest during a flood. Calculating $T(v)$ (Eq. 5) entails the determination of the expected flood duration $D(v)$ and the probabilities $\vec{F}(v)$. The calculation of these quantities requires a solution to Eq. 6 and Eq. 8. Analytical solutions for extremely simple cases were presented in Sec. 4. More complicated systems, such as for real rivers, seem analytically intractable. A computer based solution for determining $D(v)$ and $\vec{F}(v)$ is required.

To determine $D(v)$, $\vec{F}(v)$ must be calculated first. Two approaches for calculating $\vec{F}(v)$ are available. One approach is to solve the ordinary differential equation (ODE) in Eq. 8 as a boundary value problem. Press et al. [10] presents two options for solving ODE boundary value problems, (1) the shooting method, and (2) relaxation methods. We implemented the shooting method for some simple systems, and the algorithm is numerically intensive as well as unstable. The instability is due to the singularities near volumes v such that $\lambda_i - q(v) = 0$. Relaxation methods, which approximate Eq. 8 by finite difference equations and attempt to minimize the error of the approximation, seem to overcome this difficulty. We are still in the process of implementing an effective algorithm for large river systems.

When solving with the relaxation method, the following initializations are required. The boundary conditions are $\vec{F}(v_{\max}) = \vec{\pi}^T$ and $\vec{F}(v_{\min}) = 0$, where $v_{\max} = q^{-1}(\lambda_{\max})$, $v_{\min} = q^{-1}(\lambda_{\min})$ are the maximum and minimum possible water levels, and $\vec{\pi}$ is the steady state probability for the Markov chain, with $\vec{\pi} M = 0$.

A second approach for calculating $\vec{F}(v)$ it to integrate the partial differential equation in Eq. 6 forward in time until stable values for $\vec{F}(v)$ are achieved. Arbitrary initial conditions will converge to the same stable values, assuming the stability of the numerical methods and the irreducibility of the Markov chain for the state of the tributaries. The ODE relaxation method is preferred, as it is computationally less intensive.

The remaining quantity needed for calculating the return time of floods of level v is the expected flood duration $D(v)$. This may be determined by numerical simulation of Eq. 6 through time. The simulation supposes that a flood of volume v has started at time $t = 0$, and keeps track of the probability that the flood recedes below v at each time $t > 0$. From this information, the expected flood duration is calculated as

$$D(v) = \int_{t=0}^{\infty} t \, \mathrm{Prob}(\text{flood recedes at time } t) \, dt \tag{13}$$

For the simulation, the flood values and time steps are discretized. This turns Eq. 6 into a finite difference equation, and turns the integration in Eq. 13 into a sum. No existing algorithms seem ready to solve this equation without some modification. We are presently implementing a simulation based on ideas for PDE numerical methods found in Press et al. [10].

The initial conditions for the simulation use the fact that for a flood of volume v to start, the tributaries must be in a state s such that the rate of water coming in is greater than the rate of discharge, $\lambda_s > q(v)$. The probability of being in state s, then, given that a flood of volume v has just begun, is $F_s(v)/(\sum_{i|\lambda_i > q(v)} F_i(v))$. This leads to the following initial conditions for a simulation to determine $D(v)$.

$$F_s(v';t=0) = \begin{cases} \frac{F_s(v)}{\sum_{i|\lambda_i > q(v)} F_i(v)} & \lambda_s > q(v), v' > v + \Delta v \\ \frac{(v'-v)F_s(v)}{\Delta v \sum_{i|\lambda_i > q(v)} F_i(v)} & \lambda_s > q(v), v' \in (v, v+\Delta v) \\ 0 & \lambda_s > q(v), v' < v \\ 0 & \lambda_s \leq q(v) \end{cases} \tag{14}$$

for some negligibly small $\Delta v > 0$. This simulation calculates $D(v)$ for only a single value of v. For multiple values of $D(v)$, the present formulation requires multiple simulations.

There are asymptotic results for evaluating $D(v)$. Since $D(v)$ is a decreasing function of v, these results may be used to bound the return period for a flood. Suppose Chézy's equation, $q(v) = av^b$, has power $b > 0$, and $\lambda_{\max}$ is the discharge for only one state, $s_{\max}$. Then the expected flood duration is asymptotically the expected time spent in state $s_{\max}$. Formally,

$$\lim_{v \to v_{\max}} D(v) = 1/\mu_{s_{\max}, s_{\max}}.$$

For the case $q(v) = av^0 = a$, the flood duration is asymptotically constant, $\lim_{v\to\infty} D(v) = \text{const}$, where the constant depends on the discharge a and the MMFS, M and Λ.

6 Discussion

The design of flood defense systems requires a statistical model to describe the frequency of extreme floods. There is an increasing awareness that these models should reflect the hydrological information about the river catchments they describe.

Our approach to finding the expected frequency for extreme floods differs in several ways from the more traditional statistical fits to annual maxima (AM) and peaks over threshold (POT) data. For AM and POT data, a line is fit to data points plotted on special paper, such as log paper. These lines have been used to infer the relation between river level and flood return frequency.

In our approach, a continuous time model for the river dynamics is used to determine the frequency of extreme floods. Our sample calculations show the importance of incorporating river dynamics by illustrating the effects of Chézy's equation parameters on the shape of the curves relating the river level and flood return frequency. These shapes do not always conform to the curves found for traditional models. In particular, the relation is not necessarily linear on log paper, as with the Gumbel model.

For one example shown here, an increase in the power parameter b of Chézy's equation led to a non-linear relation on log paper. As b increased, the slope of the curve relating flood volume and the frequency of extreme floods decreased. We hypothesize that this may be true for large v for more complicated systems as well. If this hypothesis is true, then flood protection designs based on drawing straight lines on log paper would be conservative for extremely rare floods. The exploration of this hypothesis analytically and/or numerically constitutes future research.

Acknowledgments We wish to thank the supporters of the Engineering Probability: Damage Accumulation conference in Delft, Holland, which made the presentation of earlier versions of this work [11] possible.

References

[1] H. E. J. Berger, *Flow Forecasting for the River Meuse.* PhD thesis, Technical University of Delft, Dept. of Civil Engineering, 1992.

[2] U.S. Water Resources Council, *Guidelines for Determining Flood Flow Frequency.* Reston, VA: Bulletin 17B of the Hydrology Committee, US Government Printing Office, 1982.

[3] J. H. A. Wijbenga *et al.*, "Toetsing uitsangspunten rivierdijkversterkingen." Deelrapport 2, Maatgevende belastingen, Ministrie van Verkeer en Waterstaat, januari 1993. (in Dutch).

[4] P.H.A.J.M. Van Gelder, "De Hoogwaterramp van Limburg in December 1993, een statistische beschouwing." Rijkswaterstaat, Dienst Wegen Waterbouwkunde, 7 March 1994. (in Dutch).

[5] C. Cunnane, "Review of statistical models for flood frequency estimation," in *Hydrologic Frequency Modeling* (V. P. Singh, ed.), (Dordrecht), pp. 49–95, International Symposium on Flood Frequency Analyses, D. Reidel Pub. Co., 1987.

[6] D. Mitra and A. I. Elwalid, "Statistical multiplexing with loss priorities in rate-based congestion control of high-speed networks," *IEEE Transactions on Communications*, vol. 42, pp. 2989–3002, Nov. 1994.

[7] S. M. Ross, *Stochastic Processes.* New York: John Wiley & Sons, Inc., 1983.

[8] K. Petersen, *Ergodic Theory.* Cambridge: Cambridge University Press, 1983.

[9] S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer.* Redwood City, CA: Addison-Wesley Pub. Co., Inc., 2nd ed., 1991.

[10] W. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C.* Cambridge: Cambridge University Press, 2nd ed., 1992.

[11] S. E. Chick, J. Shortle, P. van Gelder, and M. B. Mendel, "A physics-based approach to predicting the frequency of extreme river levels," in *Engineering Probability: Flood Protection* (R. M. Cooke and M. B. Mendel, eds.), Amsterdam: Kluwer, to appear, 1996.

A Summary of Notation

$B(z, a, b)$	incomplete beta function
$D(v)$	expected duration of flood of volume v or greater (time)
δ_i	lag of water flow from tributary i to river cross-section (time)
$F_s(v; t)$	Prob$\{V \leq v,$ tributary state $S(t) = s;$ at time t$\}$
$\vec{F}(v; t)$	vector of $F_i(v; t)$
$\vec{F}(v)$	steady state probabilities for river system
h	height of river (m)
$I(z, a, b)$	incomplete regularized beta function
$q(v)$	discharge from river cross-section (m^3/sec)
$q(v) = av^b$	Chézy's equation for discharge (m^3/sec)
$\bar{q}$	maximal discharge from tributaries (m^3/sec)
K	number of major tributaries which affect river cross-section
$M = \mu_{ij}$	Markov transition matrix for tributary state changes
N	number of states ($\mid \mathcal{L} \mid^K$) of tributaries
$\vec{\pi}$	stationary probability tributary states
l_i	water input rate due to tributary i (m^3/sec)
λ_s	discretized river input when tributaries in state s (m^3/sec)
Λ	diag(λ_s)
$\mathcal{L}$	discrete set of discharge values for tributaries
$S(t)$	state of tributaries, $S(t) = s \in \mathcal{L}^K$ at time t
$T(v)$	expected return time for flood of volume v or greater (time)
V, v	volume of water in a river cross-section (m^3)

B Appendix

Throughout, suppose that N, Λ, M, and $\vec{\pi}$ are as specified in Sec. 4.

Proof of Eq. 5 To show that $T(v) = D(v)/P(v)$, results of ergodic theory are used. Technical assumptions used include (1) the river states are described by an irreducible Markov chain (see, eg [7]), (2) the Markov chain is finite, (3) derivatives exist so that Eq. 6 is meaningful. More general arguments might be made to relax the second assumption.

First, using assumption (2), it is equivalent to analyze the system with a uniformized version of the process, with time step h. Second, the following arguments shows that ergodic theory applies to the river system described by the Markov chain state and volume, (s, v). This is true by virtue of the (a) ergodicity of irreducible Markov chains, (b) the fact that for the river system considered here, adding the volume of water of the cross-section does not change the irreducibility, and (c) the fact that Eq. 6 conserves probability. A result of Halmos (see, eg [8], Ch. 2, Thm. 3.2) insures that the process is ergodic.

From the mean ergodic theorem, the return time of an event A is $1/\text{Prob}(A)$. In our case, the event of interest is $A = \{V_t \geq v\} \cap \{V_{t-h} < v\}$. (The fact that if (s_t, v_t) is ergodic, then $(s_t, v_t, s_{t-h}, v_{t-h})$ is ergodic is implicitly used). This probability is $P(v)/D_h(v)$, where the probability $P(v)$ of exceeding v is not changed by the uniformization, and $D_h(v)$ is the expected duration of the flood when flood durations are rounded up to the next multiple of h (again, a result of the mean ergodic theorem). As $h \to 0$, $D_h(v) \to D(v)$, as required.

Proof of Eq. 9. Set $g(v) = \mu_{12}F_1(v) - \mu_{21}F_2(v)$. Eq. 8 can be rewritten

$$g' = -\left(\frac{\mu_{12}}{\lambda_1 - cv} + \frac{\mu_{21}}{\lambda_2 - cv}\right) g$$

This has solution

$$g(v) = k_1(\lambda_1 - cv)^{\frac{\mu_{12}}{c}}(cv - \lambda_2)^{\frac{\mu_{21}}{c}}$$

where k_1 is some constant. By definition of $g(v)$,

$$F_1'(v) = \frac{-1}{\lambda_1 - cv} g(v) \text{ and } F_2'(v) = \frac{1}{\lambda_2 - cv} g(v)$$

Integrating shows that the F_i are multiples of an incomplete beta function with the parameters given in Eq. 9. The boundary condition $F_1(\lambda_1/c) = \pi_1 = \mu_{21}/(\mu_{12} + \mu_{21})$ determines the constant of proportionality.

Proof of Eq. 10 The expected length of a flood $D(v - \Delta v)$ is found in terms of $D(v)$. It is assumed that Δv is small and that $v \in (\lambda_2/a, \lambda_1/a)$. With high probability, the length of a $(v - \Delta v)$ flood is the same as the length of a level v flood *plus* the extra time the river is in the region $[v - \Delta v, v]$ (flood B in Figure

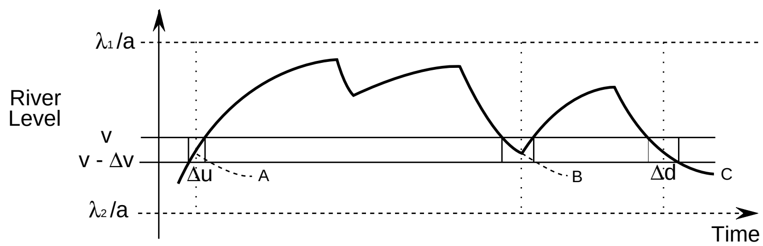

Figure 9: A sample flood of height $(v - \Delta v)$.

9). These extra up and down crossing times are given by the river flow equation, $\dot{v} = \lambda_i - cv$:

$$\Delta u = \frac{\Delta v}{\lambda_1 - av} + o(\Delta v^2), \qquad \Delta d = \frac{\Delta v}{av - \lambda_2} + o(\Delta v^2). \tag{15}$$

It is also possible that the river up-turns in the region $[v - \Delta v, v]$, generating another level v flood (flood C). Another possibility is that the river down-turns before it reaches height v (flood A).

These scenarios are summarized mathematically as follows. The probability that the river does not down-turn in the first time interval Δu is $(1 - \mu_{12}\Delta u) + o(\Delta v^2)$. If there is no down-turn, then there are a series of level v floods, where the number of such floods is a geometric random variable with mean $1/p = (1 + \mu_{21}\Delta d) + o(\Delta v^2)$. The expected length of each such flood is $(\Delta u/2 + D(v) + \Delta d/2) + o(\Delta v^2)$, where the expected extra time spent in $[v - \Delta v, v]$ is included. Finally, the time at the very beginning and end of the flood is added, $(\Delta u + \Delta d)/2$. This gives the following equation *to first order* in Δv:[2]

$$D(v - \Delta v) = (1 - \mu_{12}\Delta u)(1 + \mu_{21}\Delta d)\left(D(v) + \frac{\Delta u + \Delta d}{2}\right) + \frac{\Delta u + \Delta d}{2}. \tag{16}$$

Substituting Eq. 15 into Eq. 16 and taking the limit as $\Delta v \to 0$ gives:

$$\frac{dE(v)}{dv} + \left(\frac{\mu_{21}}{av - \lambda_2} - \frac{\mu_{12}}{\lambda_1 - av}\right) E(v) = -\left(\frac{1}{av - \lambda_2} + \frac{1}{\lambda_1 - av}\right). \tag{17}$$

Eq. 10 is the solution to this first-order linear differential equation (where $\psi(\lambda_1/c)$ is the constant of integration). It can be checked by L'Hopital's rule

[2] All scenarios which contribute second order terms to this equation are not considered. For instance, the probability of an up-turn *and* a down-turn while the river remains in $[v - \Delta v, v]$ is second order.

that Eq. 10 satisfies the desired boundary conditions – namely, $D(\lambda_2/a) = \infty$ and $D(\lambda_1/a) = 1/\mu_{12}$.

Proof of Eq. 11 The proof follows the same strategy as for Eq. 9. The only difference is that there is no bound on the maximal volume of water, so the boundary conditions are $F_1(\infty) = \mu_{21}/(\mu_{21} + \mu_{12})$, $F_2(\infty) = \mu_{12}/(\mu_{21} + \mu_{12})$. Further, when it is raining it is impossible to have an empty buffer, so $F_1(0) = 0$. These determine all constants of integration.

Proof of Eq. 12 Along similar lines, if the discharge is $q(v) = d$ when $v > 0$, the analog of Equation 17 is

$$\frac{dD(v)}{dv} + \left(\frac{\mu_{21}}{d-\lambda_2} - \frac{\mu_{12}}{\lambda_1 - d}\right) D(v) = -\left(\frac{1}{d-\lambda_2} + \frac{1}{\lambda_1 - d}\right). \qquad (18)$$

The general solution to this first-order linear differential equation with constant coefficients is

$$D(v) = \frac{\frac{1}{d-\lambda_2} + \frac{1}{\lambda_1-d}}{\frac{\mu_{21}}{d-\lambda_2} - \frac{\mu_{12}}{\lambda_1-d}} + k_2 \exp\left[\left(\frac{\mu_{12}}{\lambda_1-d} - \frac{\mu_{21}}{d-\lambda_2}\right) v\right],$$

for some constant k_2. This equation can be simplified to

$$D(v) = \frac{\bar{\lambda} - \lambda_2}{\mu_{21}(d-\bar{\lambda})} + k_2 \exp\left[\left(\frac{\mu_{12}}{\lambda_1-d} - \frac{\mu_{21}}{d-\lambda_2}\right) v\right],$$

It remains to calculate k_2. The fraction of time that $v = 0$ is $F_2(0) = (d-\bar{\lambda})/(d-\lambda_2)$, by the steady state equations. The expected time the system is empty is $1/\mu_{21}$, so by the renewal/reward theorem

$$F_2(0) = \frac{1/\mu_{21}}{D(0^+) + 1/\mu_{21}}$$

Solving for $D(0^+)$ in terms of $F_2(0)$ and using the steady state solution,

$$D(0^+) = \frac{\bar{\lambda} - \lambda_2}{\mu_{21}(d-\bar{\lambda})}$$

Therefore, $k_2 = 0$.

The implication is that all floods of volume $v > 0$ have the same expected duration. This makes sense, given the memoryless properties of the tributary states and the constant input/discharge rates.

List of Figures