Open Access

Advances in Environmental and Engineering Research

LIDSEN Publishing Inc.

Research Article

Estimation of the Return Periods in Hydrology

Mario Lefebvre *

Department of Mathematics and Industrial Engineering, Polytechnique Montréal, C.P. 6079, Succursale Centre-ville, Montréal H3C 3A7, Canada; E-Mail: mlefebvre@polymtl.ca

* **Correspondence:** Mario Lefebvre; E-Mail: mlefebvre@polymtl.ca

Academic Editor: Wen-Cheng Liu

Special Issue: Advances in Hydrology, Water Quality and Sediment Simulation Modelling

Adv Environ Eng Res
2022, volume 3, issue 4
doi:10.21926/aeer.2204049

Received: September 26, 2022
Accepted: November 29, 2022
Published: December 08, 2022

Abstract

A filtered Poisson process is proposed as a model for river flows. With the help of real-life data, the model parameters are estimated. Mathematical formulae are derived in order to estimate the various return periods of the river. An application to two rivers shows that the point estimates are very close to the corresponding values computed by hydrologists, based on historical data. Moreover, by modifying the values of the parameters in the model, we can see the potential effects of climate change on the return periods.

Keywords

Filtered Poisson process; point estimate; climate change

1. Introduction

Let $X(t)$ be the flow of a certain river at time t. We consider the following model:

$$X(t) = \sum_{n=1}^{N(t)} w\,(t, \tau_n, Y_n) \quad (X(t) = 0 \text{ if } N(t) = 0), \tag{1}$$

in which the τ_n's are the arrival times of the events of a Poisson process $\{N(t), t \geq 0\}$ with rate λ. Moreover, $Y_1, Y_2, \ldots$ are independent and identically distributed random variables, and are also independent of $\{N(t), t \geq 0\}$. The stochastic process $\{X(t), t \geq 0\}$ is known as a *filtered Poisson process*; see Parzen [1].

In hydrology, the *response function* w is often taken to be of the form

$$w(t, \tau_n, Y_n) = Y_n\, e^{-(t-\tau_n)/c},$$

and the Y_i's are assumed to be exponentially distributed with parameter μ. Thus, λ is the rate at which signals occur, $1/\mu$ is the average size of a signal, and c is a parameter that is related to the speed at which the effect of a signal on the flow discharge decreases. Figure 1 shows an example of a trajectory of a filtered Poisson process having the above response function.

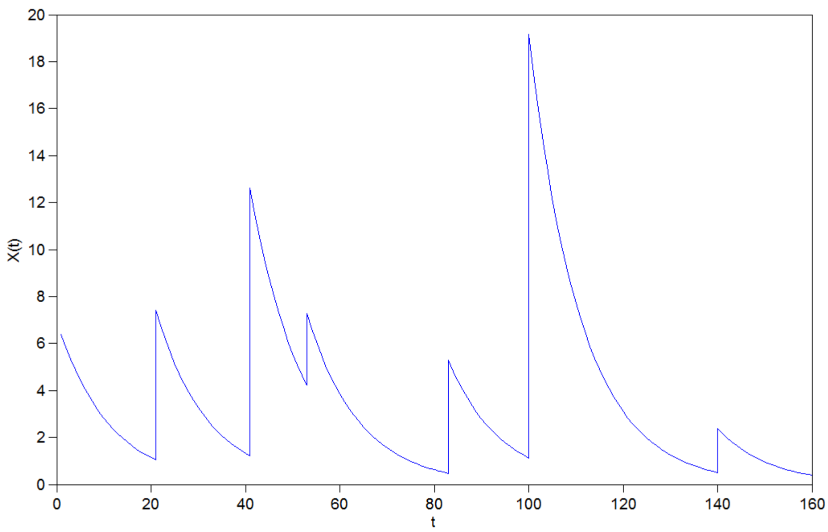

Figure 1 Example of a trajectory of a filtered Poisson process.

The author has used filtered Poisson processes (see for instance Lefebvre and Guilbault [2]) and their generalization to filtered renewal processes (see Lefebvre [3] and Lefebvre and Bensalma [4]) in various hydrological applications. In particular, they can serve as models for daily flows, but also to estimate the probability that the river flow will exceed a given threshold [5].

Another important problem in hydrology is the estimation of the return periods or recurrence intervals. A return period T_n, where n is in years, is defined such that the probability of exceedance of a certain event (like a large river discharge flow) during a given year is equal to $1/n$:

$$P[\exists\, t \in (0{,}365]: X(t) \geq T_n] = \frac{1}{n}.$$

The probability is usually expressed in percentage. The estimates of return periods are generally based on historical data over a long period of time. These return periods are important in risk analysis, in particular in the design of structures such as dams or bridges. Sometimes, the return periods are computed for the height or water level of the river, rather than its flow.

Gumbel [6] wrote a seminal paper in which he applied the theory of largest values to flood flows. He used his method to estimate the return periods of the Rhône River in France, and the Mississippi River in the USA, and he compared his estimates to the observed return periods. Since then, many papers have been published on this subject; see, for example, Onyutha and Willems [7]. Jennings *et al.* [8] wrote a report on the use of regression equations to estimate the magnitude and frequency of floods. There are also papers on return periods of extreme precipitation [9], hail storms [10], hydrological droughts [11], etc.

It is important to mention that the above-mentioned papers are based on *statistical techniques*. Akyuz *et al.* [12] used Markov chain models to estimate drought characteristics. Here, we will present a technique based on a continuous-time stochastic process to estimate return periods. This stochastic process has been shown in previous papers by the author to be an appropriate model for river flows. The technique that we propose is different from the ones that can be found in other papers and it will be seen that it yields very good results.

In computing the estimates of the T_n's, it is assumed that they do not vary over time and that they do not depend on past events. However, because of climate change, these assumptions become more and more dubious; see also Serinaldi [13].

In this note, we will present a method that could enable us to see the effects of climate change on the values of the return periods. The method will be presented in the next section, and it will be applied to two rivers in Section 3. We will conclude with a few remarks in Section 4.

2. Estimates of the Return Periods

In practice, flow values are generally recorded on a daily basis. Therefore, we have 365 observations of $X(t)$ per year. Let M denote the annual maximum flow discharge. If we assume that the various peaks during a given year are sufficiently spaced in time, then

$$P[M \leq m] \simeq \prod_{i=1}^{365\lambda} P\left[Y_i \leq m\right] \overset{i.i.d.}{=} (1 - e^{-\mu m})^{365\lambda}. \quad (2)$$

Let p denote the probability of having two events on two consecutive days:

$$p := P[\tau_k - \tau_{k-1} \leq 1] = P[\mathrm{Exp}(\lambda) \leq 1] = 1 - e^{-\lambda}.$$

The probability of having r events on r consecutive days is therefore, by independence, p^{r-1}.

Remark. By looking at real-life hydrographs, it is actually not obvious to determine whether a peak was caused by a single event, or by multiple consecutive events.

If λ is large enough, and if we neglect the possibility of observing three or more consecutive signals, a more precise formula for the distribution function of M would be

$$P[M \leq m] \simeq (1 - e^{-\mu m})^{365\lambda(1-p)} \left\{P\left[Y_1 e^{-1/c} + Y_2 \leq m\right]\right\}^{365\lambda(\mathrm{p}/2)}. \quad (3)$$

Indeed, there are, on average, $365\lambda(1-p)$ single events and $365\lambda p/2$ pairs of consecutive events. Moreover, we assumed that

$$Y_1 \leq Y_1 e^{-1/c} + Y_2,$$

which should be the case in general since Y_1 and Y_2 have the same exponential distribution.

If we use Eq. (2) to approximate the distribution function $F_M(m)$ of M, we need to estimate the parameters λ and μ, while with Eq. (3) we must estimate c as well. If the value of λ is small, then so is p, so that one can make use of the simpler formula (2) to approximate $F_M(m)$.

We can write that

$$F_M(T_n) = 1 - \frac{1}{n}.$$

Therefore, if we have the values of T_n, for various n, that were calculated by hydrologists, we can use them to estimate the parameters in our model.

In the next section, we will present an application to the Delaware River, which is an important river located in the United States, and to the Lim River in Montenegro. The technique that we propose can be applied to any river. However, in order to evaluate its accuracy, we need some estimates calculated by hydrologists to compare our point estimates to theirs. We could also, in theory, compare the point estimates derived from our mathematical formulae to the corresponding ones calculated by making use of the statistical techniques that can be found in the papers mentioned in Section 1. However, because the values provided by hydrologists are assumed to be reliable and are used in practice, it is preferable to compare our point estimates to these estimates.

3. Applications

The values of various return periods have been estimated for the Delaware River; see Schopp and Firda [14]. The estimated values (in cubic feet per second) at the Montague, N.J., station are presented in Table 1.

Table 1 Estimated return periods for the Delaware River.

n	2	5	10	25	50	100	500
T_n	62500	101000	127000	164000	194000	226000	308000

First, we will estimate the parameters λ and μ in Eq. (2). To do so, we need two values of T_n. Because the estimates are most likely more reliable for small values of n, we chose $n = 2$ and $n = 5$. We have

$$F_M(T_2) = 0.5 \quad \Leftrightarrow \quad (1 - e^{-62500\mu})^{365\lambda} = 0.5$$

and

$$F_M(T_5) = 0.8 \quad \Leftrightarrow \quad (1 - e^{-101000\mu})^{365\lambda} = 0.8.$$

We find that the point estimate of μ is $\hat{\mu} \simeq 0.0000278$, which implies that $\hat{\lambda} \simeq 0.0098$. Hence, the average size of flow increases due to (important) signals is 35971 f^3/s, and there are, on average, 3.577 such signals per year.

Next, we compute

$$\hat{p} \simeq 1 - e^{-0.0098} \simeq 0.00975.$$

Therefore, in this application we can neglect the possibility that there will be r signals on r consecutive days, for $r = 2,3, \ldots$

With $\hat{\mu}$ and $\hat{\lambda}$, we can estimate the values of T_n for any n, based on our model; see Table 2. We see that the model underestimates the value of T_n for n very large. However, estimating a flow that occurs on average every 500 years is quite difficult and requires a lot of observations. Therefore, the point estimate T_{500} provided by hydrologists is probably more or less reliable, and a confidence interval for T_{500} should be very wide. Onyutha and Willems [7] wrote that in practice return periods between 5 and 100 years are used for the design of hydraulic structures, while higher return periods around T_{100} are used mainly for flood plain development and medium-sized flood protection works. Moreover, T_{500} is rarely used in designs.

Table 2 Point estimates $\hat{T}_n$ of T_n based on the model for the Delaware River.

n	2	5	10	25	50	100	500
$\hat{T}_n$	62500	101000	127000	161200	186350	211450	269500
T_n	62500	101000	127000	164000	194000	226000	308000

Now, the aim of this work is to try to forecast the effects of climate change on the return periods. To do so, we computed the new point estimates $\hat{T}_n$ if firstly the parameter μ is replaced by 0.9μ (so that $\frac{1}{\mu}$ becomes $\frac{10}{9\mu}$), secondly λ is increased to 1.1λ, and finally when both changes are made.

The results are presented in Table 3.

Table 3 Point estimates $\hat{T}_n$ for the Delaware River when the parameters are modified.

n	T_n	$\hat{T}_n$	0.9μ	1.1λ	Both
2	62500	62500	69500	65600	72900
5	101000	101000	112300	104250	116000
10	127000	127000	141550	131000	145500
25	164000	161200	179100	164550	182800
50	194000	186350	207100	189800	211000
100	226000	211450	235000	215000	239000
500	308000	269500	299500	272800	303300

We see that the value of μ has more influence on the return period than that of λ. Moreover, in the case of μ, the value of $\hat{T}_n$ is almost exactly multiplied by $10/9$ for each n, while the percentage increase of $\hat{T}_n$ decreases with n when λ is replaced by 1.1λ.

Finally, we will apply the technique that we propose to estimate the return periods of the Lim River, located in Montenegro. The return periods (in m^3/s) based on the years 1961-1990 at the station Bijelo Polje can be found in a report (p. 159) on the Second National Communication on Climate Change [15] and are given in Table 4. In the report, there are also forecasts for these return periods in the periods 2001-2030 and 2071-2100. Notice that T_n is given from $n = 5$ to $n = 10000$, which is an extremely large value of n. Moreover, in cubic feet per second, $T_5 \simeq 19741$, compared to 101000 for the Delaware River. Therefore, the Lim River is a mid-size river in comparison to the Delaware.

Table 4 Estimated return periods for the Lim River.

n	5	10	20	25	40	50	100	200	500	1000	2000	10000
T_n	559	678	792	828	903	939	1049	1159	1303	1413	1522	1776

As above, we first estimate the parameters λ and μ in Eq. (2), using T_5 and T_{10}. We have

$$F_M(T_5) = 0.8 \quad \Leftrightarrow \quad (1 - e^{-559\mu})^{365\lambda} = 0.8$$

and

$$F_M(T_{10}) = 0.9 \quad \Leftrightarrow \quad (1 - e^{-678\mu})^{365\lambda} = 0.9.$$

This time, we find that $\hat{\mu} \simeq 0.00625$ and $\hat{\lambda} \simeq 0.0198$. The average size of flow increases is 160 m^3/s, and there are 7.227 important signals per year.

Because

$$\hat{p} \simeq 1 - e^{-0.0198} \simeq 0.0196,$$

the possibility that there will be r signals on r consecutive days is again negligible for $r = 2,3, \ldots$

In Table 5, we give the estimated values of T_n based on our model. The values of $\hat{T}_n$ are very close to the corresponding T_n's, for n from 5 to 10000. The difference in absolute value between $\hat{T}_{10000}$ and T_{10000} is less than 1%. We see that $\hat{T}_n$ very slightly overestimates T_n from $n = 25$. As in the case of the Delaware River, we could easily compute the new values of $\hat{T}_n$ when we modify the parameters μ and λ.

Table 5 Point estimates $\hat{T}_n$ of T_n based on the model for the Lim River.

n	5	10	20	25	40	50	100	200	500	1000	2000	10000
$\hat{T}_n$	559	678	792	829	905	941	1053	1164	1311	1422	1533	1790
T_n	559	678	792	828	903	939	1049	1159	1303	1413	1522	1776

4. Conclusion

In this note, we proposed a model for the flow of a river that enables us to estimate the return periods for that river. In two applications, we saw that, after having estimated the parameters in our model, we obtained good or very good estimates of the various return periods, when compared to the values provided by hydrologists. Then, we computed the new point estimates for the Delaware River when the values of the parameters are modified, as we can expect, because of climate change.

As a sequel to this work, we could try to apply the same technique to estimate the return periods of other important rivers, to see if the results that we obtained are as robust as they appear to be.

To do so, we must of course have the corresponding return periods estimated by hydrologists. Finally, we could also try to use different models for the flow $X(t)$, in particular models based on diffusion processes.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). The author also wishes to thank the anonymous reviewers of this paper for their constructive comments.

Author Contributions

The author did all the research work of this study.

Competing Interests

The author reports that there are no competing interests to declare.

References

1. Parzen E. Stochastic processes. Philadelphia: Society for Industrial and Applied Mathematics; 1962. 324p.
2. Lefebvre M, Guilbault JL. Using filtered Poisson processes to model a river flow. Appl Math Model. 2008; 32: 2792-2805.
3. Lefebvre M. A filtered renewal process as a model for a river flow. Math Probl Eng. 2005; 2005: 49-59.
4. Lefebvre M, Bensalma F. An application of filtered renewal processes in hydrology. Int J Eng Math. 2014; 2014: 593243.
5. Lefebvre M, Bensalma F. On the distribution of a river flow at the time of the next increase. Can J Civ Eng. 2015; 42: 1146-1148.
6. Gumbel EJ. The return period of flood flows. Ann Math Stat. 1941; 12: 163-190.
7. Onyutha C, Willems P. Empirical statistical characterization and regionalization of amplitude–duration–frequency curves for extreme peak flows in the Lake Victoria Basin, East Africa. Hydrol Sci J. 2015; 60: 997-1012.
8. Jennings ME, Thomas Jr WO, Riggs HC. Nationwide summary of us geological survey regional regression equations for estimating magnitude and frequency of floods for Ungaged Sites, 1993. Reston, VA: USGS Publications Warehouse; 1994; Report 94-4002.
9. DeGaetano AT. Time-dependent changes in extreme-precipitation return-period amounts in the continental United States. J Appl Meteorol Climatol. 2009; 48: 2086-2099.
10. Fraile R, Berthet C, Dessens J, Sánchez JL. Return periods of severe hailfalls computed from hailpad data. Atmos Res. 2003; 67: 189-202.
11. Sharma TC, Panu US. Predicting return periods of hydrological droughts using the Pearson 3 distribution: A case from rivers in the Canadian prairies. Hydrol Sci J. 2015; 60: 1783-1796.
12. Akyuz DE, Bayazit M, Onoz B. Markov chain models for hydrological drought characteristics. J Hydrometeorol. 2012; 13: 298-309.
13. Serinaldi F. Dismissing return periods! Stoch Environ Res Risk Assess. 2015; 29: 1179-1189.

14. Schopp RD, Firda GD. Flood magnitude and frequency of the Delaware River in New Jersey, New York, and Pennsylvania. Reston, VA: U.S. Geological Survey; 2008; Open-File Report 2008-1203.
15. UNFCCC. The second national communication on climate change of Montenegro to UNFCCC [Internet]. Bonn: The United Nations Framework Convention on Climate Change; 2015. Available from: https://unfccc.int/sites/default/files/resource/mnenc2_eng.pdf.