

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/282844456>

Hundred years of return period: Strengths and limitations

Article in *Water Resources Research* · October 2015

DOI: 10.1002/2015WR017820

CITATIONS

62

READS

916

5 authors, including:



Elena Volpi

Università Degli Studi Roma Tre

72 PUBLICATIONS 2,303 CITATIONS

[SEE PROFILE](#)



Aldo Fiori

Università Degli Studi Roma Tre

189 PUBLICATIONS 4,219 CITATIONS

[SEE PROFILE](#)



Salvatore Grimaldi

Tuscia University

166 PUBLICATIONS 5,403 CITATIONS

[SEE PROFILE](#)



Federico Lombardo

Ministero dell'Interno

60 PUBLICATIONS 751 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



The uncertainty of atmospheric processes in planning a hybrid renewable energy system for a non-connected island [View project](#)



Investigation of the stochastic nature of solar radiation for renewable resources management [View project](#)

100 years of Return Period: Strengths and limitations

E. Volpi,¹ A. Fiori,¹ S. Grimaldi,²³ F. Lombardo,¹ and D. Koutsoyiannis⁴

Corresponding author: E. Volpi, Department of Engineering, University of Roma Tre, Via Vito
Volterra, 62, 00146 Roma, Italy (elena.volpi@uniroma3.it)

¹Department of Engineering, University of
Roma Tre, Via V. Volterra, 62, 00146
Rome, Italy

²Department for Innovation in Biological,
Agro-food and Forest systems (DIBAF),
University of Tuscia, Via San Camillo De
Lellis snc, 01100 Viterbo, Italy

³Honors Center of Italian Universities
(H2CU), Sapienza University of Rome, Via
Eudossiana 18, 00184 Roma, Italy

⁴Department of Water Resources, Faculty
of Civil Engineering, National Technical
University of Athens, Heroon Polytechniou
5, 15780 Zographou, Greece

1 **Abstract.** 100 years from its original definition by *Fuller* [1914], the prob-
2 abilistic concept of return period is widely used in hydrology as well as in
3 other disciplines of geosciences to give an indication on critical event rareness.
4 This concept gains its popularity, especially in engineering practice for de-
5 sign and risk assessment, due to its ease of use and understanding; however,
6 return period relies on some basic assumptions that should be satisfied for
7 a correct application of this statistical tool. Indeed, conventional frequency
8 analysis in hydrology is performed by assuming as necessary conditions that
9 extreme events arise from a stationary distribution and are independent of
10 one another. The main objective of this paper is to investigate the proper-
11 ties of return period when the independence condition is omitted; hence, we
12 explore how the different definitions of return period available in literature
13 affect results of frequency analysis for processes correlated in time. We demon-
14 strate that, for stationary processes, the independence condition is not nec-
15 essary in order to apply the classical equation of return period (i.e. the in-
16 verse of exceedance probability). On the other hand, we show that the time-
17 correlation structure of hydrological processes modifies the shape of the dis-
18 tribution function of which the return period represents the first moment.
19 This implies that, in the context of time-dependent processes, the return pe-
20 riod might not represent an exhaustive measure of the probability of failure,
21 and that its blind application could lead to misleading results. To overcome
22 this problem, we introduce the concept of Equivalent Return Period, which

²³ controls the probability of failure still preserving the virtue of effectively com-
²⁴ municating the event rareness.

1. Introduction

25 “The storm event had a return period of 30 years” or “this dam spillway was designed
26 for a 1000-year return period discharge” are two classical statements that one could read
27 or hear everyday. High-school students could read them in newspapers, housewives could
28 hear them at the market or hydrologists could write them in a technical report. This simple
29 example recalls that the return period is the most ubiquitous statistical concept adopted
30 in hydrology but also in many other disciplines (seismology, oceanography, geology, etc...).

31 It appears that the concept of return period was first introduced by *Fuller [1914]* who
32 pioneered statistical flood frequency analysis in the USA. Return period finds wide pop-
33 ularity mainly because it is a simple statistical tool taken from engineering practices
34 [*Gumbel, 1958*]. For example, engineers who work on flood control are interested in the
35 expected time interval at which an event of given magnitude is exceeded for the first time,
36 which gives a definition of the return period. Another common definition is the average of
37 the time intervals between two exceedances of a given threshold of river discharge. From
38 a logical standpoint, the first definition is as justifiable as the second one; they generally
39 differ, even though they become practically indistinguishable if consecutive events are
40 independent in time. Both are used in hydrology [*Fernández and Salas, 1999a, b*] and, in
41 this paper, we will show how they may affect the frequency analysis applications under
42 certain conditions.

43 The return period is inversely related to the probability of exceedance of a specific
44 value of the variable under consideration (e.g. river discharge). For example, the annual
45 maximum flood-flow exceeded with a 1% probability in any year is called the 100-year

46 flood. Therefore, a T -year return period does not mean that one and only one T -year
47 event should occur every T years, but rather that the probability of the T -year flood being
48 exceeded is $1/T$ in every year [*Stedinger et al.*, 1993].

49 The traditional methods for determining the return period of extreme hydrologic events
50 assume as key conditions that extreme events (i) arise from a stationary distribution, and
51 (ii) are independent of one another. The hypotheses of stationarity and independence
52 are commonly assumed as necessary conditions to proceed with conventional frequency
53 analysis in hydrology [*Chow et al.*, 1988]. Recently, the former assumption has been
54 questioned by several researchers [e.g. *Cooley*, 2013; *Salas and Obeysekera*, 2014; *Du*
55 *et al.*, 2015; *Read and Vogel*, 2015]. However, we endorse herein the following important
56 statement by *Gumbel* [1941] about the general validity of stationarity assumption. “In
57 order to apply any theory we have to suppose that the data are homogeneous, i.e. that no
58 systematical change of climate and no important change in the basin have occurred within
59 the observation period and that no such changes will take place in the period for which
60 extrapolations are made. It is only under these obvious conditions that forecasts can be
61 made”. The reader is also referred to *Koutsoyiannis and Montanari* [2015] and *Montanari*
62 *and Koutsoyiannis* [2014], where it can be noted that many have lately questioned the
63 stationarity assumption, but careful investigation of claims made would reveal that they
64 mostly arise from the confusion of dependence in time with nonstationarity.

65 The purpose of this paper is to investigate the properties of return period when the
66 independence condition is omitted. In hydrology, indeed, dependence has been recognized
67 by many scientists to be the rule rather than the exception since a long time [e.g. *Hurst*,
68 1951; *Mandelbrot and Wallis*, 1968]. The concept of dependence in extreme events relates

69 to the fact that the occurrence of a high or low value for the variable of interest (e.g.
70 river discharge) has some influence on the value of any succeeding observation. *Leadbetter*
71 [1983] found that the type of the limiting distribution for maxima is unaltered for weakly
72 dependent occurrences of extreme events. We demonstrate that, under general depen-
73 dence conditions, the classical relationship between the return period and the exceedance
74 probability is again unaltered. On the other hand, we investigate the impact of the de-
75 pendence structure on the shape of the distribution function of which the return period
76 represents the first moment.

77 Based on the papers by *Fernández and Salas* [1999a], *Sen* [1999], and *Douglas et al.*
78 [2002] we first summarize in Section 2 the available definitions of return periods (aver-
79 age occurrence interval - and - average recurrence interval) specifying the mass function
80 equations and the related return period formulae. Moreover, in Section 2.2 and 2.3 the
81 independent and time-dependent cases are analyzed in detail, while an Appendix provides
82 the proof that the widely used return period equation (average recurrence interval) is not
83 affected by the dependence structure of the process of interest. However, in Section 2.3
84 it is pointed out that the time-dependence influences the shape of the interarrival time
85 distribution function and the probability of failure.

86 Two illustrative examples, i.e. using a two-state Markov process and an autoregressive
87 process, are described in Section 3 and results are discussed in Section 4 in order to in-
88 vestigate further the theoretical premises depicted in Sections 2.2 and 2.3. Besides, to
89 overcome the difficulties that arise from the application of the return period concept in a
90 time-dependent context, we propose in Section 4.1 the adoption of an Equivalent Return
91 Period (*ERP*), which resembles the classical definition of return period in the case of in-

92 dependence while it is able to control the probability of failure under the time-dependence
93 condition. The ERP can be useful to avoid introducing the concept of probability of fail-
94 ure in engineering practice. Indeed, the latter may not be as simple to understand as the
95 return period, which is a well-established concept in applications, routinely employed by
96 practitioners.

97 Concluding remarks discuss the obtained results by stressing caution against using the
98 concept of return period blindly given that multiple definitions exist. However, we confirm
99 the virtue of return period showing that the classical formulation is insensitive to the time-
100 dependence condition.

2. Return period and probability of failure

2.1. Mathematical framework

101 Let $Z(\tau)$ be a stochastic process that characterizes a natural process typically evolving
102 in continuous time τ . As observations of $Z(\tau)$ are only made in discrete time, it is assumed
103 here that the observations are made at constant time intervals $\Delta\tau$, and this interval is
104 considered the unit of time. Hence, we consider the corresponding discrete-time process
105 that is obtained by sampling $Z(\tau)$ at spacing $\Delta\tau$, i.e. $Z_j = Z(j\Delta\tau)$ where j ($= 1, 2, \dots$)
106 denotes discrete time. For convenience, herein we express discrete time as $t = j - j_0$,
107 where j_0 is the current time step; therefore the discrete-time process is indicated as Z_t
108 and $t = 0$ denotes the present. We assume that Z_t is a stationary process [*Papoulis*, 1991];
109 thus, it is fully described up to the second order properties by its marginal probability
110 function and its autocorrelation structure. Generally, in this paper we use upper case
111 letters for random variables or events, and lower case letters for values, parameters, or
112 constants.

113 We are interested in the occurrence of possible excursions of Z_t above/below a high/low
 114 level (threshold) z , which may determine the failure of a structure or system. In particular,
 115 we define a dangerous event as $A = \{Z > z\}$, which is an extreme maximum; anyway,
 116 A could be any type of extreme event, i.e. maximum or minimum. In the following we
 117 denote by p the probability of the event $B = \{Z \leq z\}$, which is the complement of A ; the
 118 probability of the event A is given by $1 - p = \Pr\{Z > z\} = \Pr A$.

119 In hydrological applications, it is usually assumed that the event A will occur on average
 120 once every return period T , where T is a time interval and, for annual observations (i.e.,
 121 $\Delta\tau = 1$ year), a number of years. In other words, the average time until the threshold z
 122 is exceeded equals T years [*Stedinger et al.*, 1993], such as

$$\frac{T}{\Delta\tau} = E[X] = \sum_{t=1}^{\infty} t f_X(t) \quad (1)$$

123 where X is the number of discrete time steps to the occurrence of an event A , $f_X(t) =$
 124 $\Pr\{X = t\}$ is its probability mass function (pmf) and $E[\cdot]$ denotes expectation. The
 125 definition of the return period leads to the formulation of the so-called *probability of*
 126 *failure* $R(l)$ (also known in literature as "risk", even if it does not account for damages)
 127 which measures the probability that the event A occurs at least once over a specified
 128 period of time: the design life l (e.g. in years) of a system or structure, where $l/\Delta\tau$ is a
 129 positive integer. Mathematically, we have

$$R(l) = \Pr\{X \leq l/\Delta\tau\} = \sum_{t=1}^{l/\Delta\tau} f_X(t) \quad (2)$$

130 Thus, the probability of failure is nothing else than the distribution function $F_X(t)$ com-
 131 puted at $t = l/\Delta\tau$.

132 As mentioned in the Introduction, two different definitions of the return period are
 133 available in the hydrological literature [see, e.g., *Fernández and Salas, 1999a* and *Douglas*
 134 *et al., 2002*]. The return period T may be defined as:

- 135 (i) the mean time interval required to the *first occurrence* of the event A ,
- 136 (ii) the mean time interval between *any two successive occurrences* of the event A .

137 Definition (i) assumes that an event A occurred in the past (at $t < 0$); the discrete time
 138 elapsed since the last event A to the current time step $t = 0$ is defined as *elapsing time* and
 139 it is denoted here as t_e ; the sketch in Figure 1 illustrates the variables used in the present
 140 analysis. In this work, we assume that time t_e can be either deterministically known or
 141 unknown and investigate implications of both conditions on the analytical formulation of
 142 the return period. Under definition (i), the return period is based on the *waiting time*
 143 (W), i.e. the number of time steps between $t = 0$ and the next occurrence of A (see
 144 Figure 1). The sum of the waiting time and the elapsing time is denoted as *interarrival*
 145 *time* $N = W + t_e$.

146 If we assume that t_e is unknown, the probability mass function of the waiting time
 147 is given by the joint probability of the sequence of events $(B_1, B_2, ..B_{t-1}, A_t)$ (see, e.g.,
 148 *Fernández and Salas, 1999a*)

$$f_W(t) = \Pr(B_1, B_2, ..B_{t-1}, A_t) \quad (3)$$

149 where A_t (B_t) is the event A (B) occurred at time t . Instead, if t_e is determin-
 150 istically known, the pmf of the waiting time is given by the joint probability of
 151 the sequence of events $(B_1, B_2, ..B_{t-1}, A_t)$ conditioned to the realization of the events
 152 $(A_{-t_e}, B_{-t_e+1}, ..B_{-1}, B_0)$ occurred at $t \leq 0$, i.e.

$$\begin{aligned}
f_{W|t_e}(t) &= \Pr(B_1, B_2, \dots, B_{t-1}, A_t | A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0) \\
&= \frac{\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0, B_1, B_2, \dots, B_{t-1}, A_t)}{\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0)}
\end{aligned} \tag{4}$$

153 Definition (ii) assumes that an event A has just occurred at $t = 0$. In such a case $t_e = 0$
154 and the waiting time W is identical to the interarrival time N . The pmf of the interarrival
155 time f_N is therefore a special case of equation (4), for $t_e = 0$, i.e.

$$\begin{aligned}
f_N(t) &= \Pr(B_1, B_2, \dots, B_{t-1}, A_t | A_0) \\
&= \frac{\Pr(A_0, B_1, B_2, \dots, B_{t-1}, A_t)}{\Pr A_0}
\end{aligned} \tag{5}$$

156 Note that Figure 1 depicts a more general case than the one represented by equation (5).
157 In the Figure, we assume that two successive occurrences of the dangerous event A are
158 at times $-t_e$ and t . Then, N is the time elapsed between the two. As stated above, the
159 specific case expressed by equation (5) can be obtained by setting $t_e = 0$. Moreover, we
160 stress here that the relation $N = W + t_e$ in the Figure holds only in the case the elapsing
161 time t_e is known, i.e. when we account for the conditional waiting time $W|t_e$.

162 It is interesting to note that the probability distributions of the unconditional (W , equa-
163 tion (3)) and conditional ($W|t_e$, equation (4)) waiting time are interrelated. In Appendix
164 A we derive some useful relations between the return periods T_W , $T_{W|t_e}$ and T_N .

165 Substituting f_W (equation (3)), $f_{W|t_e}$ (equation (4)) or f_N (equation (5)) to f_X in (1)
166 and (2), we obtain the expressions of the return periods T_W , $T_{W|t_e}$ and T_N and of the
167 corresponding probabilities of failure R_W , $R_{W|t_e}$ and R_N , respectively. In general, the
168 probability mass functions given by equations (3) to (5) are expected to have different

169 shapes, leading to different values of the return period of the event A . In the following,
 170 we illustrate and discuss the differences among the above definitions when varying the
 171 correlation structure of the process Z_t ; specifically, we study first the independent case,
 172 which is customary in hydrological applications, and then the more general case with some
 173 positive correlation in time (persistent case).

2.2. Independent case

174 If Z_t is a purely random process, then its random variables are mutually independent
 175 and their joint probability distribution equals the product of marginal ones. Therefore, we
 176 may write e.g. $\Pr(B_0, B_1, B_2, \dots, B_{t-1}, A_t) = \Pr B_0 \Pr B_1 \dots \Pr A_t$. Substituting in equations
 177 (3), (4) and (5) the products of the marginal exceedance or non-exceedance probabilities
 178 and thanks to the stationarity assumption (that implies $\Pr A_t = 1 - p$ and $\Pr B_t = p$
 179 for any t), we can derive the same geometric distribution in all cases. Therefore, $f_W =$
 180 $f_{W|t_e} = f_N = f$, with

$$f(t) = p^{t-1} (1 - p) \quad (6)$$

181 It follows from equation (1) that the return period T ($T = T_W = T_{W|t_e} = T_N$) is given
 182 by

$$\frac{T}{\Delta\tau} = \frac{1}{1 - p} \quad (7)$$

183 while the variance of the pmf (6) is $v = p/(1 - p)^2$. From equation (6), it also follows
 184 that the probability of failure given by equation (2) becomes

$$R(l) = 1 - \left(1 - \frac{\Delta\tau}{T}\right)^{l/\Delta\tau} = 1 - p^{l/\Delta\tau} \quad (8)$$

185 where again $R = R_W = R_{W|t_e} = R_N$.

186 Thus, for the independent case all the definitions of return period collapse to the same
 187 expression (7). This result, which is well known in the literature [e.g. *Stedinger et al.*,
 188 1993], builds on the fact that in the independent case the occurrence of an event at any
 189 time $t \leq 0$ does not influence what happens afterwards.

2.3. Persistent case

190 Although independence of Z_t is usually invoked for the derivation of equation (7) [e.g.
 191 *Kottegoda and Rosso*, 1997, p. 190], it is possible to show that the mean interarrival time
 192 T_N is equal to (7) also in case of processes correlated in time; the general proof, which
 193 is given here for the first time, is illustrated in detail in Appendix B. The same property
 194 was shown by *Lloyd* [1970] for the particular case of a Markov chain process. As shown
 195 in Appendix B, equation (7) for the mean interarrival time holds true, regardless of the
 196 type of the correlation structure of Z_t .

197 Even though the dependence structure of the process Z_t does not affect the expected
 198 value of N (i.e., T_N), we show that this is not the case with its pmf f_N (see equation
 199 (5)). Let us consider a process characterized by a positive correlation in time. If a
 200 dangerous event A occurs at $t = 0$, then the conditional probability of occurrence of
 201 another dangerous event at $t = 1$ will be greater than $1 - p$ (independent case); this
 202 yields that the probability mass function $f_N(t)$ will have a larger mass for $t = 1$ and a
 203 lower mass elsewhere with respect to the independent case (equation (6)). Hence, while
 204 the mean value remains the same, the variance of the interarrival time N is larger than

205 that of the independent case and it increases with the temporal correlation. This implies
 206 that the probability of failure R_N (following equation (2)) is strongly affected by the
 207 time-dependence structure of the process.

208 Conversely, the return periods T_W and $T_{W|t_e}$ do account for the temporal correlation of
 209 Z_t . Recalling that $(1 - p) = 1/\mathbb{E}[N]$ (see equations (7) and (1)), it can be shown that
 210 (see Appendix A, equation (A8))

$$\frac{T_W}{T_N} = \frac{1}{2} \left(\frac{\mathbb{E}[N^2]}{\mathbb{E}[N]^2} + \frac{1}{\mathbb{E}[N]} \right) \quad (9)$$

211 Equation (9) shows that T_W is greater than or equal to T_N . It is easy to check that
 212 $T_W = T_N$ for independent processes, in line with the discussion reported under Section
 213 2.2. When the process is correlated in time, the term $\mathbb{E}[N^2]/\mathbb{E}[N]^2$ is expected to increase
 214 with the autocorrelation of the process, thus resulting in the inequality $T_W > T_N$. Hence,
 215 the mean waiting time is generally larger than the mean interarrival time for temporally
 216 correlated processes.

217 In the following Sections we will examine the pmfs of the waiting times W and $W|t_e$ and
 218 the interarrival time N , as well as their average values (T_W , $T_{W|t_e}$ and T_N), as functions
 219 of the temporal correlation of the process. To this end, we make use of two different
 220 illustrative examples, the first is based on a Markov chain, while the second uses an
 221 AR(1) model. For convenience - and without loss of generality - $\Delta\tau$ is set equal to one.

3. Illustrative examples

3.1. Example 1: two state Markov-dependent process

222 We consider here a stochastic process Z_t which is based on a Markov chain Y_t . This
 223 process is considered here since it allows to easily derive the analytical expressions of

224 the probability mass functions of the waiting and interarrival times, as done in previous
 225 literature works by *Lloyd* [1970], *Rosbjerg* [1977] and *Fernández and Salas* [1999a]. The
 226 Markov chain Y_t has two states, which here represent the events $A_t = \{Z_t > z\}$ and
 227 $B_t = \{Z_t \leq z\}$ with probability $1 - p$ and p , respectively. For the Markov property, the
 228 probability of a state at a given time t depends solely on the state at the previous time
 229 step $t - 1$, e.g. $\Pr(B_t|B_{t-1}...B_0) = \Pr(B_t|B_{t-1})$. Applying the chain rule to the Markov
 230 property (e.g. *Papoulis*, 1991, p. 636), it follows that the joint probability of a sequence
 231 of states, e.g. $\Pr(B_1, B_2, ..B_t) = \Pr\{Z_1 \leq z, Z_2 \leq z, \dots, Z_t \leq z\}$, can be written as
 232 $\Pr(B_1) \Pr(B_2|B_1) \dots \Pr(B_t|B_{t-1}) = \Pr\{Z_1 \leq z\} \Pr\{Z_2 \leq z|Z_1 \leq z\} \dots \Pr\{Z_t \leq z|Z_{t-1} \leq$
 233 $z\}$.

234 The process Z_t described above is indicated in the following as two state Markov-
 235 dependent process and denoted by 2Mp. For each value of p (i.e. of z) Z_t is fully charac-
 236 terized by the marginal probabilities of the states A and B ($1 - p$ and p) and by the tran-
 237 sition probability matrix, $M = [[\Pr(A_{t+1}|A_t), \Pr(A_{t+1}|B_t)], [\Pr(B_{t+1}|A_t), \Pr(B_{t+1}|B_t)]]$
 238 where $\Pr(A_{t+1}|A_t) + \Pr(B_{t+1}|A_t) = 1$ and $\Pr(A_{t+1}|B_t) + \Pr(B_{t+1}|B_t) = 1$. We denote
 239 by q the joint probability of non-exceedance of the threshold value z for two successive
 240 events, i.e. $q = \Pr(B_{t+1}, B_t)$ for any t ; it ensues that $M = [[1 - (p - q) / (1 - p), 1 -$
 241 $q/p], [(p - q) / (1 - p), q/p]]$.

242 The probability mass function of the unconditional waiting time f_W (equation (3))
 243 becomes

$$f_W(t) = \begin{cases} 1 - p & (t = 1) \\ p \left(\frac{q}{p}\right)^{t-2} \left(1 - \frac{q}{p}\right) & (t \geq 2) \end{cases} \quad (10)$$

244 with mean given by

$$T_W = 1 + \frac{p^2}{(p - q)} \quad (11)$$

245 and variance $\text{var}[W] = p^2(p - p^2 + q)/(p - q)^2$. After substituting equation (10) in (2),
 246 the probability of failure in a period of length l is given by

$$R_W(l) = 1 - p \left(\frac{q}{p}\right)^{l-1} \quad (12)$$

247 while the pmf of the conditional waiting time $f_{W|t_e}$ (equation (4)) for $t_e > 0$ reduces to

$$f_{W|t_e}(t) = \left(\frac{q}{p}\right)^{t-1} \left(1 - \frac{q}{p}\right) \quad (13)$$

248 with mean

$$T_{W|t_e} = \frac{p}{(p - q)} \quad (14)$$

249 and variance $\text{var}[W|t_e] = pq/(p - q)^2$. The probability of failure based on the conditional
 250 waiting time is given by

$$R_{W|t_e}(l) = 1 - \left(\frac{q}{p}\right)^{l-1} \quad (15)$$

251 Equation (14) shows how for the 2Mp model the mean waiting time distribution is not
 252 affected by the value of t_e . This builds upon the fact that the conditional non-exceedance
 253 probability at t depends only on that at $t - 1$, due to the property of the Markov chain.

254 Finally, the pmf of the interarrival time N (equation (5)) assumes the following expres-
 255 sion

$$f_N(t) = \begin{cases} 1 - (p - q)/(1 - p) & (t = 1) \\ \frac{(p-q)}{(1-p)} \left(\frac{q}{p}\right)^{t-2} \left(1 - \frac{q}{p}\right) & (t \geq 2) \end{cases} \quad (16)$$

256 while its mean is given by equation (7) with $\Delta\tau = 1$ (following the general proof given in
 257 Appendix B), and the variance is equal to $\text{var}[N] = p(p - 2p^2 + q) / [(p - 1)^2(p - q)]$. The
 258 probability of failure in a period of length l is given by

$$R_N(l) = 1 - \frac{p - q}{1 - p} \left(\frac{q}{p}\right)^{l-1} \quad (17)$$

259 The joint probability q may assume values in the range $[\max(2p - 1, 0), p]$: the lower
 260 and upper bounds correspond to perfect negative and positive correlations in time, respec-
 261 tively; in the independent case, $q = p^2$. We consider here only processes positively corre-
 262 lated (i.e. persistent), as it is commonly the case in hydrology (e.g. rainfall and discharge);
 263 thus, $q \in [p^2, p]$. Furthermore, we assume that Z_t is a standard Gaussian process and that
 264 the joint probability q is ruled by a bivariate Gaussian distribution; under the latter
 265 assumption, q can be described in terms of the lag-1 autocorrelation coefficient ρ . Specifi-
 266 cally, q is computed as $q = \Pr\{Z_{t+1} \leq z, Z_t \leq z\} = \int_{-\infty}^z \int_{-\infty}^z f_{\mathbf{Z}}(z_t z_{t+1}; \mathbf{0}, \Sigma_2) dz_{t+1} dz_t$
 267 where $f_{\mathbf{Z}}$ is the probability density function of the bivariate Gaussian distribution
 268 $\mathcal{N}_2(\mathbf{Z}; \mathbf{0}, \Sigma_2)$ with zero mean and $\Sigma_2 = \{\{1, \rho\}, \{\rho, 1\}\}$, with $\rho \in [0, 1]$. Note that ρ
 269 denotes the correlation in the parent process Z_t and not that between the events exceed-
 270 ing the threshold, i.e. $A = \{Z > z\}$. The correlation between the extremes is ruled by
 271 the shape of the parent bivariate distribution, which is assumed here to be Gaussian; the
 272 latter assumption implies that the correlation between the events A is negligible to null
 273 for high threshold values, since the Gaussian process is asymptotic independent.

3.2. Example 2: AR(1) process

274 We now assume that Z_t follows an AR(1) process (first-order autoregressive process), i.e.
 275 $Z_t = \rho Z_{t-1} + \alpha_t$ where ρ is the lag-1 correlation coefficient and $\alpha_t \sim \mathcal{N}\left(0, \sqrt{1 - \rho^2}\right)$, such
 276 that the process is characterized by a multivariate Gaussian distributions $\mathcal{N}_t(\mathbf{Z}; \mathbf{0}, \mathbf{\Sigma}_t)$
 277 with $\mathbf{Z} = \{Z_1, Z_2..Z_t\}$ and $\mathbf{\Sigma}_t = \{\rho^{|i-k|}\}$, $i, k = 1..t$. We assume again $\rho \in [0, 1]$.

278 Even if conceptually simple and similar to the 2Mp (see e.g. *Saldarriaga and Yevjevich,*
 279 1970), AR(1) is rather different in terms of the pmfs f_W , $f_{W|t_e}$ and f_N . Both the processes
 280 are based on the Markov property; however, in AR(1) the Markov property applies to the
 281 continuous random variable Z and not to the state $Y = \{Z \leq z\}$. It means that in AR(1)
 282 the joint probability $f_{\mathbf{Z}}(z_1, z_2..z_t)$ can be expressed as $f_Z(z_1)f_Z(z_2|z_1)..f_Z(z_t|z_{t-1})$, while
 283 the same simplification cannot apply to the joint probability of a sequence of states, e.g.
 284 $\Pr(B_1, B_2, ..B_t) = \Pr\{Z_1 \leq z, Z_2 \leq z, \dots, Z_t \leq z\}$, as for 2Mp. The joint probability of
 285 any sequence can be estimated by proper integration of the joint pdf of the multivariate
 286 Gaussian distribution \mathcal{N}_t . This entails that the pmfs f_W , $f_{W|t_e}$ and f_N , given by equations
 287 (3), (4) and (5) respectively, cannot be simplified as in the case of 2Mp, but they can be
 288 written as

$$f_W(t) = \int_{-\infty}^z \int_{-\infty}^z \dots \int_z^{+\infty} f_{\mathbf{Z}}(z_1, z_2..z_t; \mathbf{0}, \mathbf{\Sigma}_t) dz_1 dz_2 .. dz_t \quad (18)$$

$$f_{W|t_e}(t) = \frac{\int_z^{+\infty} \int_{-\infty}^z \dots \int_z^{+\infty} f_{\mathbf{Z}}(z_{-t_e}, z_{-t_e+1}..z_t; \mathbf{0}, \mathbf{\Sigma}_{t+t_e}) dz_{-t_e} dz_{-t_e+1} .. dz_t}{\int_z^{+\infty} \int_{-\infty}^z \dots \int_{-\infty}^z f_{\mathbf{Z}}(z_{-t_e}, z_{-t_e+1}..z_0; \mathbf{0}, \mathbf{\Sigma}_{t_e}) dz_{-t_e} .. dz_0} \quad (19)$$

289 while f_N can be derived from the latter under the assumption $t_e = 0$. Finally, substituting
 290 the previous expressions in (1) and (2) we get the corresponding return periods and
 291 probabilities of failure.

292 Interestingly enough, unlike the 2Mp, $f_{W|t_e}$ (19) depends on t_e , i.e. the elapsing
 293 time. This relies on the fact that the conditional non-exceedance probability at t , i.e.
 294 $\Pr(B_t|B_{t-1}\dots B_0)$, generally depends on the whole sequence of previous events for AR(1),
 295 while it only depends on that at $t - 1$ for the 2Mp. In such a sense, AR(1) is more
 296 correlated than 2Mp.

4. Results and discussion

297 We start this Section by discussing the effects of temporal correlation on the probability
 298 mass functions f_W (equation (10)) and f_N (equation (16)), and the related return periods
 299 T_W , T_N (equation (11) and (7) with $\Delta\tau = 1$, respectively) for the two state Markov-
 300 dependent process (2Mp).

301 Figure 2 illustrates T_W and T_N as functions of the independent return period T (i.e. of
 302 the non-exceedance marginal probability p) for several values of the correlation coefficient
 303 ρ . It is seen that T_N equals T , being independent of ρ as demonstrated in Appendix B;
 304 for $\rho = 0$ (black line) it is always $T_N = T_W = T = (1 - p)^{-1}$. Conversely, the mean
 305 waiting time T_W increases with ρ (equation (9)); T_W is always greater than the mean
 306 interarrival time T_N , which thus represents a lower bound for the return period (Figure
 307 2a). Specifically, for values of T around 5, T_W is roughly eight times larger than T_N for
 308 $\rho = 0.99$ and about twice for $\rho = 0.75$; for small and very large values of T (i.e. for
 309 small and high values of the threshold z , respectively) T_W tends to the independent limit
 310 $T = (1 - p)^{-1}$ (Figure 2b).

311 As discussed in Section 2.3, although $T_N = T$ for any ρ , the pmf f_N (as well as f_W) may
 312 be significantly influenced by the correlation structure of the Z_t process. The distribution
 313 functions of W and N are illustrated in Figure 3, for various values of ρ and $p = 0.9$.

314 The mean values for each distribution (i.e. the return periods normalized with respect to
 315 $\Delta\tau = 1$) are denoted by the vertical dashed lines. The broadness of both distributions
 316 increases with ρ , as also indicated by the increase of their variance and skewness (not
 317 shown).

318 Figure 3a shows that the distribution function computed at T_W , which corresponds to
 319 the probability of failure in the period T_W (see equation (2)), is independent of ρ taking
 320 approximately the value 0.63 for high values of p [*Stedinger et al.*, 1993].

321 On the other hand, F_N changes dramatically when increasing temporal correlation ρ .
 322 This may result in very high values of the probability of failure for the same T_N , even
 323 for small time intervals t (Figure 3b). Thus, although the return period T_N remains the
 324 same for correlated and independent processes (all the vertical dashed lines corresponding
 325 to the different values of ρ collapse into a unique line, depicted in black), the probability
 326 that the threshold z is exceeded in the period T_N can be much larger for the former than
 327 for the latter (up to about 0.9 for the limit case $\rho = 0.99$).

328 We now illustrate and discuss the probability functions for W , $W|t_e$ and N for the AR(1)
 329 process, as well as the corresponding mean values, as functions of the lag-1 autocorrelation
 330 coefficient ρ . Results are compared to those obtained for the previously analyzed 2Mp
 331 case.

332 The probability mass functions f_W (equations (18)) and f_N (equation (19) for $t_e = 0$)
 333 for AR(1) are similar to those for 2Mp, even if they are characterized by a much larger
 334 dispersion, and thus they are not shown here. Their averages T_W and T_N are depicted
 335 in Figure 4, as function of the independent return period T , for $\rho = 0.75$ and $\rho = 0.99$.
 336 T_W and T_N for AR(1) (continuous lines) are also compared to those pertaining to the 2Mp

337 (dashed lines). The mean waiting times T_W for the two models are similar, although T_W
 338 is generally larger for AR(1); since the two processes have the same ρ , this result is a
 339 direct consequence of the stronger correlation of AR(1) with respect to 2Mp, as explained
 340 in previous Section. Larger differences are expected for even more persistent processes,
 341 i.e. processes characterized by a longer range persistence with respect to the AR(1).

342 As mentioned in the previous Section, the stronger correlation of AR(1) also influences
 343 the mean conditional waiting time $T_{W|t_e}$, which depends on the elapsing time t_e in contrast
 344 to that of 2Mp. $T_{W|t_e}(t_e)$ is illustrated in Figure 5 for $p = 0.9$ and for a few values of the
 345 correlation coefficient ρ . For each value of ρ , $T_{W|t_e}$ is by definition equal to the mean inter-
 346 arrival time T_N for $t_e = 0$ (see equation (4)); $T_{W|t_e}$ increases with t_e tending to an asymp-
 347 totic value that is greater than T_W (dashed lines). This behaviour arises from the fact that
 348 the conditional non-exceedance probability $(B_1, B_2, \dots, B_{t-1}, A_t | A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0)$ (eq.
 349 4) depends on the whole sequence of previous events. However, as t_e becomes very high
 350 the previous dangerous event A_{-t_e} has occurred too distant in time to significantly affect
 351 the realization of the next event at time t ; the latter is mainly controlled by a sequence
 352 of antecedent events whose length strictly depends on the shape of the autocorrelation
 353 function of the underlying process Z_t . Due to the exponential shape of the AR(1) auto-
 354 correlation function, i.e. $\rho_t(t) = \rho^t$, $T_{W|t_e}$ is expected to approach the asymptotic value
 355 when t_e becomes larger than the integral scale of the process, $\lambda(\rho) = 1/(1 - \rho)$.

356 Conversely, $T_{W|t_e}$ for 2Mp maintains a constant value for any $t_e > 0$ since the con-
 357 ditional joint probability in equation (4) $\Pr(B_1, B_2, \dots, B_{t-1}, A_t | A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0)$ de-
 358 pends only on the state at $t = 0$, due to the Markov property of the Y_t chain (as already

discussed in Section 3.1); moreover, being influenced by a longer sequence of safe events
 (B), both $T_{W|t_e}$ and T_W of AR(1) are larger than those of 2Mp (results not shown).

We finally explore how the probabilities of failure $R_W(T_W)$, $R_{W|t_e}(T_{W|t_e})$ and $R_N(T_N)$
 behave as functions of the correlation coefficient ρ ; results are summarized in Figure 6 for
 the processes 2Mp and AR(1) and compared to the independent case. For both processes,
 the probability of failure based on the interarrival time (N) may assume values much
 larger than the independent case; $R_N(T_N)$ significantly increases with the autocorrelation
 of the process ρ , (compare e.g. 2Mp for $\rho = 0.75$ and $\rho = 0.99$) and, more generally, with
 the correlation structure of the process (compare AR(1) and 2Mp for the same value of ρ).
 On the contrary, when we consider the waiting time W (conditional and unconditional),
 the probability of failure is less than the independent case. This reduction is significant
 when we account for the elapsing time t_e , thus when we add information about the last
 dangerous event occurred in the past. Note that Figure 6 specifically refers to the cases
 $t_e = 10$ for AR(1) while it is representative of any $t_e > 0$ for 2Mp. As for AR(1),
 $R_{W|t_e}(T_{W|t_e})$ reduces with respect to the independent case when t_e is much larger than
 the integral scale of the process, i.e. $t_e > \lambda$ when $\rho = 0.75$ ($\lambda = 4$) (Fig. 6a); conversely,
 when the event A has happened in the recent past (when $\rho = 0.99$, we have $t_e < \lambda$ with
 $\lambda = 100$), the conditional waiting time for high p has a behaviour which approaches that
 of the interarrival time (i.e. with higher probability of failure than the independent case,
 as in Figure 6b).

4.1. Equivalent Return Period (ERP)

The return period is a means of expressing the exceedance probability. Despite being
 a standard term in engineering applications (in engineering hydrology in particular), the

381 concept of return period is not always an adequate measure of the probability of failure
382 and has been sometimes incorrectly understood and misused [*Serinaldi*, 2014]. The results
383 discussed in previous Section strengthen the above message, extending it to correlated Z_t
384 fields (with Markovian dependence); for the cases examined here, the statistics of the
385 waiting or interarrival time show negligible differences with respect to the independent
386 case for small values of ρ , while they are strongly affected by the autocorrelation when
387 $\rho \gtrsim 0.5$ (see Figures 2 and 5). Consequently, using directly the probability of failure
388 in engineering practice could be a better choice under the latter condition. However,
389 although more effective and appropriate, the probability of failure may not be as simple
390 to understand as the return period, which is already an established concept in applications
391 and routinely employed by practitioners.

392 To overcome this problem, we introduce the concept of "equivalent" return period
393 (*ERP*). Its aim is to retain the relative simplicity of the return period concept and
394 extend it to temporally correlated hydrological variables; for correlated processes, *ERP*
395 is defined to be the period that would lead to the same probability of failure pertaining to
396 a given return period T in the framework of classical statistics (independent case). Hence,
397 *ERP* resembles the classical definition of return period in the case of independence, thus
398 preserving its simplicity and strength in indicating the event rareness; in addition it is
399 able to control the probability of failure under the time-dependence condition.

400 *ERP* can be defined starting from the concept of interarrival time (N) or waiting
401 time (W). Practitioners should adopt the most appropriate definition according to the
402 circumstances, the task and the data available. If the time t_e elapsed since the last
403 dangerous event is known, it could be adopted the definition based on the conditional

404 waiting time, or that based on the interarrival time in the case $t_e = 0$; the latter could
 405 be the case where an existing structure failed because of an event A and the immediate
 406 construction of another structure is needed (as discussed by *Fernández and Salas* [1999a]).

407 In the case we are accounting for the interarrival time (N), ERP can be calculated
 408 assuming $R_N(ERP) = R(T)$ where R_N is the probability of failure based on the inter-
 409 arrival time (equation (2) for $f_X = f_N$), while $R(T)$ is given by equation (8) for $l = T$.
 410 For the 2Mp R_N is given by equation (17) (where $\Delta\tau = 1$) when $l = ERP$; thus, the
 411 analytical formulation of ERP can be easily derived as

$$ERP = 1 + \frac{\ln \frac{1-p}{p-q} + \frac{1}{1-p} \ln p}{\ln \frac{q}{p}} \quad (20)$$

412 For the AR(1), R_N can be numerically computed by substituting equation (19) in (2).

413 In the case of more complex models for the simulation of hydrological quantities, ERP
 414 could be computed directly by numerical Monte Carlo simulations.

415 Figure 7 depicts the behaviour of ERP as function of T , for both the AR(1) (continuous
 416 lines) and 2Mp processes (dashed lines; equation (20) with $p = 1 - 1/T$). The figure
 417 shows that the values of ERP and T tend to coincide asymptotically; this is especially
 418 so for small correlation coefficients. For a given T , the value of ERP is always smaller
 419 (sometimes much smaller) than T ; differences increase with the correlation coefficient ρ
 420 and with the correlation structure of the process (compare AR(1) to 2Mp). Recalling that
 421 $T = 1/(1 - p)$, Figure 7 can be used either to determine ERP when the p -th quantile z is
 422 known (i.e., for a given event $A = \{Z > z\}$ that will be exceeded with probability $1 - p$)
 423 in risk assessment problems, or to determine the design variable (i.e. the threshold z) in
 424 terms of p once the ERP is fixed in design problems; in the latter case we choose ERP

425 and then calculate the design variable z , such that the probability of failure is equal to
426 that we should have in the independent case.

427 We emphasize that results shown here are obtained under several assumptions, such as
428 the type of temporal correlation, bivariate Gaussian distribution, etc.; this implies that, for
429 example, a different distribution may result in larger differences between the independent
430 and time-correlated conditions (due, e.g., to asymptotic dependence). Hence, further work
431 is needed to generalize the above results.

5. Conclusions

432 The return period is a critical parameter largely adopted in hydrology for risk assess-
433 ment and design. It is defined as the mean value of the waiting time to the next dangerous
434 event (T_W) or the interarrival time between successive dangerous events (T_N). As shown
435 in previous literature, both definitions lead to the same result in the case of time inde-
436 pendence of the underlying process. However, in cases of time-persistent processes the
437 two definitions lead to different expressions. Hence, we reexamine herein the above defi-
438 nitions in the context of temporally correlated processes; furthermore, by making use of
439 two illustrative examples we discuss the effects of the temporal correlation ρ of the parent
440 process on the return period and the probability of failure. The examples proposed here
441 are based on a two state Markov-dependent process (2Mp), and an AR(1) process; even
442 if the two processes share the Markov property, they are characterized by rather different
443 time distributions.

444 The main conclusions drawn in this paper are listed below.

445 • We provide a unitary framework for the estimation of the return periods T_W , T_N
446 and the related probabilities of failure R_W , R_N in the context of persistent processes:

447 we provide general relationships for the probability functions of the waiting time W (un-
448 conditional and conditional on the time t_e elapsed since the last dangerous event) and
449 the interarrival time N . The choice between W and N in applications depends on the
450 available information on past events and the type of structure.

451 • We demonstrate that the mean interarrival time T_N is not affected by the time-
452 dependence structure of the process, e.g. the correlation coefficient ρ . Thus, the well
453 known formula for independent processes is valid for any process, temporally correlated
454 or not.

455 • Although T_N is not affected by ρ , for persistent processes the corresponding proba-
456 bility of failure can be much larger than that pertaining to the independent case, which is
457 itself not negligible. Hence, the mean interarrival time T_N can easily provide a biased and
458 wrong perception of the risk of failure, especially in the presence of temporally correlated
459 hydrological variables.

460 • On the other hand, the mean waiting times effectively account for the correlation
461 structure of the hydrological process. T_W is always larger than the mean interarrival
462 time T_N , which acts as a lower bound. If the time t_e from the last dangerous event is
463 deterministically known, we can use that information to condition the waiting time W to
464 the next occurrence.

465 • The return periods T_W and $T_{W|t_e}$ typically increase with the correlation ρ . Specif-
466 ically, they depend on the overall correlation structure of the process, as highlighted by
467 comparing results for 2Mp and AR(1); in the case of processes characterized by a longer
468 range persistence with respect to the AR(1), we may expect even stronger differences.

469 • The analyses carried out here provide some further insight into the overall meaning
470 and significance of the return period, especially in view of hydrological applications, but
471 also in other geophysical fields. Despite being a simple and easy to implement metric, the
472 return period should be used with caution in the presence of time-correlated processes.
473 Indeed, the probability of failure depends on the whole shape of the probability function,
474 which in turn may strongly depend on ρ , and the return period is just the first order
475 moment; the latter may not be relevant when in presence of asymmetric and skewed
476 distributions, like e.g. some of those displayed in Figure 3.

477 • To partially overcome the above limitations, we propose to adopt in the time-
478 dependent context the Equivalent Return Period (*ERP*), which preserves the virtue of the
479 classical return period of effectively communicating the event rareness. *ERP* resembles
480 the classical definition of return period in the case of independence, while it is able to
481 control the probability of failure under the time-dependence condition.

482 We conclude with a note on the practical implications of the present analysis. Results
483 shown here highlight that the independence condition is not necessary for the application
484 of the classical return period equation; notwithstanding this, practitioners should take
485 care of the time-persistence structure of the process when estimating risk from data, to
486 correctly evaluate the probability of failure (e.g. through *ERP*). However, it is interesting
487 to stress that the differences between the correlated and uncorrelated case are small to
488 negligible when $\rho \lesssim 0.5$. Thus, the temporal correlation of the process may be safely
489 disregarded in such cases, as far as the return period is concerned.

490 **Acknowledgments.** We thank the Editor and the three anonymous Reviewers for
491 their thoughtful comments. The research has been partially funded by the Italian Min-

492 istry of University and Research through the projects PRIN 2010JHF437 and PRIN
493 20102AXKAJ. No data was used in producing this manuscript.

Appendix A: General relationships between f_W , $f_{W|t_e}$ and f_N

494 Since we can write that $\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0) = \Pr(B_0, B_1, \dots, B_{t_e-1}, A_{t_e}) =$
 495 $f_W(t_e + 1)$, the probability mass function of the conditional waiting time, $f_{W|t_e}(t)$, can
 496 be expressed as function of f_W and f_N as in the following

$$\begin{aligned}
 f_{W|t_e}(t) &= \frac{\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0, B_1, B_2, \dots, B_{t-1}, A_t)}{\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0)} & (A1) \\
 &= \frac{\Pr(B_{-t_e+1}, \dots, B_{-1}, B_0, B_1, B_2, \dots, B_{t-1}, A_t | A_{-t_e}) \Pr A_{-t_e}}{\Pr(A_{-t_e}, B_{-t_e+1}, \dots, B_{-1}, B_0)} \\
 &= \frac{(1-p)}{f_W(t_e + 1)} f_N(t + t_e)
 \end{aligned}$$

497 By making use of the simple identity $\Pr(C) = \Pr(AC) + \Pr(BC)$, which is valid for any
 498 events A and C (with B always denoting the complement of A), f_W can be expressed as
 499 function of f_N

$$\begin{aligned}
 f_W(t) &= \Pr(B_1, \dots, B_{t-1}, A_t) & (A2) \\
 &= \Pr(A_0, B_1, \dots, B_{t-1}, A_t) + \Pr(B_0, B_1, \dots, B_{t-1}, A_t) \\
 &= \Pr(B_1, \dots, B_{t-1}, A_t | A_0) \Pr A_0 + \Pr(B_0, B_1, \dots, B_{t-1}, A_t) \\
 &= f_N(t) (1-p) + f_W(t+1)
 \end{aligned}$$

500 by solving equation (A2) for f_N and substituting the resulting expression in (A1) we
 501 obtain

$$f_{W|t_e}(t) = \frac{1}{f_W(t_e + 1)} [f_W(t + t_e) - f_W(t + t_e + 1)] \quad (A3)$$

502 Since f_N is a special case of $f_{W|t_e}$, when $t_e = 0$ equation (A3)

$$f_N(t) = \frac{1}{1-p} [f_W(t) - f_W(t+1)] \quad (\text{A4})$$

503 Moreover, if we exploit the recursive property of equation (A2), we can write

$$f_W(2) = f_W(1) - (1-p)f_N(1) \quad (\text{A5})$$

$$f_W(3) = f_W(2) - (1-p)f_N(2)$$

$$= f_W(1) - (1-p)f_N(1) - (1-p)f_N(2)$$

$$f_W(4) = f_W(3) - (1-p)f_N(3)$$

$$= f_W(1) - (1-p)f_N(1) - (1-p)f_N(2) - (1-p)f_N(3)$$

...

504 thus obtaining

$$\begin{aligned} f_W(t+1) &= f_W(1) - (1-p) \sum_{k=1}^t f_N(k) & (\text{A6}) \\ &= (1-p) \left[1 - \sum_{k=1}^t f_N(k) \right] \\ &= (1-p) [1 - F_N(t)] \\ &= (1-p) \bar{F}_N(t) \end{aligned}$$

505 where we used $f_W(1) = \Pr A_1 = 1 - p$ and the survival function of N , i.e. $\bar{F}_N(t) =$

506 $1 - F_N(t) = 1 - \sum_{k=1}^t f_N(k) = \sum_{k=t+1}^{\infty} f_N(k)$. The relationship between $f_{W|t_e}$ and f_N is

507 obtained by substituting equations (A4) and (A6) into (A3)

$$f_{W|t_e}(t) = \frac{f_N(t+t_e)}{\bar{F}_N(t_e)} \quad (\text{A7})$$

508 We adopt equation (A6) to derive the analytical expression of the return period T_W as
 509 function of f_N

$$\begin{aligned}
 \frac{T_W}{\Delta\tau} &= \sum_{t=1}^{\infty} t (1-p) \bar{F}_N(t-1) = (1-p) \sum_{t=1}^{\infty} t \bar{F}_N(t-1) & (A8) \\
 &= (1-p) \sum_{t=1}^{\infty} t \sum_{k=t}^{\infty} f_N(k) = (1-p) \sum_{k=1}^{\infty} f_N(k) \sum_{t=1}^k t \\
 &= (1-p) \sum_{k=1}^{\infty} \frac{k(k+1)}{2} f_N(k) \\
 &= (1-p) \left[\sum_{k=1}^{\infty} \frac{k^2}{2} f_N(k) + \sum_{k=1}^{\infty} \frac{k}{2} f_N(k) \right] \\
 &= \frac{(1-p)}{2} (\mathbb{E}[N^2] + \mathbb{E}[N])
 \end{aligned}$$

510 Finally, substituting equation (A7) into (1) we obtain $T_{W|t_e}$ as function of f_N

$$\begin{aligned}
 \frac{T_{W|t_e}}{\Delta\tau} &= \sum_{t=1}^{\infty} t \frac{f_N(t+t_e)}{\bar{F}_N(t_e)} & (A9) \\
 &= \frac{1}{\bar{F}_N(t_e)} \sum_{t=1}^{\infty} [(t+t_e) f_N(t+t_e) - t_e f_N(t+t_e)] \\
 &= \frac{1}{\bar{F}_N(t_e)} \left[\sum_{k=t_e+1}^{\infty} k f_N(k) - t_e \sum_{k=t_e+1}^{\infty} f_N(k) \right] \\
 &= \frac{1}{\bar{F}_N(t_e)} \left[\sum_{k=t_e+1}^{\infty} k f_N(k) - t_e \bar{F}_N(t_e) \right] \\
 &= \sum_{k=t_e+1}^{\infty} \frac{k f_N(k)}{\bar{F}_N(t_e)} - t_e
 \end{aligned}$$

Appendix B: Mean interarrival time, T_N

511 Substituting equation (5), which is of general validity, in (1) we have

$$\begin{aligned}
\frac{T_N}{\Delta\tau} &= \sum_{t=1}^{\infty} t f_N(t) = 1 \Pr\{N = 1\} + 2 \Pr\{N = 2\} + \dots & (B1) \\
&= \Pr(A_1|A_0) + 2 \Pr(B_1, A_2|A_0) + 3 \Pr(B_1, B_2, A_3|A_0) + \dots \\
&= \frac{1}{\Pr A_0} [\Pr(A_0, A_1) + 2 \Pr(A_0, B_1, A_2) + 3 \Pr(A_0, B_1, B_2, A_3) + \dots] \\
&= \frac{1}{1-p} [\Pr(A_0, A_1) + 2 \Pr(A_0, B_1, A_2) + 3 \Pr(A_0, B_1, B_2, A_3) + \dots]
\end{aligned}$$

512 By making use again of the identity $\Pr(CA) = \Pr(C) - \Pr(CB)$, where B always denotes
513 the opposite event of A , we obtain

$$\begin{aligned}
\frac{T_N}{\Delta\tau} &= \frac{1}{1-p} [(\Pr A_0 - \Pr(A_0, B_1)) + 2(\Pr(A_0, B_1) - \Pr(A_0, B_1, B_2)) & (B2) \\
&\quad + 3(\Pr(A_0, B_1, B_2) - \Pr(A_0, B_1, B_2, B_3)) + \dots] \\
&= \frac{1}{1-p} [\Pr A_0 + \Pr(A_0, B_1) + \Pr(A_0, B_1, B_2) + \Pr(A_0, B_1, B_2, B_3) + \dots]
\end{aligned}$$

514 Using once more the same identity, we find

$$\begin{aligned}
\frac{T_N}{\Delta\tau} &= \frac{1}{1-p} [(1 - \Pr B_0) + (\Pr B_1 - \Pr(B_0, B_1)) & (B3) \\
&\quad + (\Pr(B_1, B_2) - \Pr(B_0, B_1, B_2)) + \dots] \\
&= \frac{1}{1-p}
\end{aligned}$$

515 which proves to be valid because of stationarity, i.e. $\Pr B_0 = \Pr B_1$, $\Pr(B_0, B_1) =$
516 $\Pr(B_1, B_2)$, etc.

References

- 517 Chow, V. T., D. R. Maidment, and L. W. Mays (1988), *Applied hydrology*, McGraw-Hill,
518 New York.
- 519 Cooley, D. (2013), Return periods and return levels under climate change, in *Extremes in*
520 *a Changing Climate*, pp. 97–114, Springer Netherlands.
- 521 Douglas, E. M., R. M. Vogel, and C. N. Kroll (2002), Impact of Streamflow Persis-
522 tence on Hydrologic Design, *Journal of Hydrologic Engineering*, 7(3), 220–227, doi:
523 10.1061/(ASCE)1084-0699(2002)7:3(220)
- 524 Du, T., L. Xiong, C. Xu, C. Gippel, S. Guo, and P. Liu (2015), Return Period and Risk
525 Analysis of Nonstationary Low-flow Series under Climate Change, *Journal of Hydrology*,
526 527, 220–227, doi:10.1016/j.jhydrol.2015.04.041.
- 527 Fernández, B., and J. D. Salas (1999a), Return period and risk of hydrologic events. II: Ap-
528 plications, *Journal of Hydrologic Engineering*, 4(4), 308–316, doi:10.1061/(ASCE)1084-
529 0699(1999)4:4(308).
- 530 Fernández, B., and J. D. Salas (1999b), Return period and risk of hydrologic events.
531 I: mathematical formulation, *Journal of Hydrologic Engineering*, 4(4), 297–307, doi:
532 10.1061/(ASCE)1084-0699(1999)4:4(297).
- 533 Fuller, W. (1914), Flood flows, *Transactions of the American Society of Civil Engineers*,
534 77, 564–617.
- 535 Gumbel, E. J. (1941), The return period of flood flows, *The annals of mathematical*
536 *statistics*, 12(2), 163–190.
- 537 Gumbel, E. J. (1958), *Statistics of Extremes*, Columbia University Press, New York.

- 538 Hurst, H. E. (1951), Long term storage capacities of reservoirs, *Transactions of the Amer-*
539 *ican Society of Civil Engineers*, 116(776-808).
- 540 Kottegoda, N. T., and R. Rosso (1997), *Probability, statistics, and reliability for civil and*
541 *environmental engineers*, McGraw-Hill, Milan.
- 542 Koutsoyiannis, D., and A. Montanari (2015), Negligent killing of scientific con-
543 cepts: the stationarity case, *Hydrological Sciences Journal*, 60(7-8), 2–22, doi:
544 10.1080/02626667.2014.959959.
- 545 Leadbetter, M. R. (1983), Extremes and local dependence in stationary sequences, *Prob-*
546 *ability Theory and Related Fields*, 65(2), 291–306.
- 547 Lloyd, E. H. (1970), Return periods in the presence of persistence, *Journal of Hydrology*,
548 10(3), 291–298.
- 549 Mandelbrot, B. B., and J. R. Wallis (1968), Noah, Joseph and operational hydrology,
550 *Water Resources Research*, 4(5), 909–918.
- 551 Montanari, A., and D. Koutsoyiannis (2014), Modeling and mitigating natural haz-
552 ards: Stationarity is immortal!, *Water Resources Research*, 50, 9748–9756, doi:
553 10.1002/2014WR016092.
- 554 Papoulis, A. (1991), *Probability, Random Variables and Stochastic Processes*, McGraw-
555 Hill, New York.
- 556 Read, L. K., and R. M. Vogel (2015), Reliability, Return Periods, and Risk under Non-
557 stationarity, *Water Resources Research*, doi:10.1002/2015WR017089.Accepted.
- 558 Rosbjerg, D. (1977), Crossing and Extremes in Dependent Annual Series, *Nordic Hydrol-*
559 *ogy*, 8, 257–266.

- 560 Salas, J. D., and J. Obeysekera (2014), Revisiting the Concepts of Return Period and
561 Risk for Nonstationary Hydrologic Extreme Events, *Journal of Hydrologic Engineering*,
562 *19*(3), 554–568, doi:10.1061/(ASCE)HE.1943-5584.0000820.
- 563 Saldarriaga, J., and V. Yevjevich (1970), *Application of run-lengths to hydrologic series*,
564 *Hydrology Paper N.40*, Colorado State University, Fort Collins.
- 565 Sen, Z. (1999), Simple risk calculations in dependent hydrological series, *Hydrological*
566 *Sciences Journal*, *44*(6), 871–878, doi:10.1080/02626669909492286.
- 567 Serinaldi, F. (2014), Dismissing return periods!, *Stochastic Environmental Research and*
568 *Risk Assessment*, pp. 1–11, doi:10.1007/s00477-014-0916-1.
- 569 Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), Frequency analysis of
570 extreme events, in *Handbook of Hydrology*, edited by D. Maidment, chap. 18, McGraw-
571 Hill, New York.

572 **List of Figures**

573 1. Illustrative sketch of the quantities involved in the definitions of the return period:
 574 excursions of the Z_t process above/below a threshold level z defining the dangerous (A_t)
 575 and safe (B_t) events.

576 2. Two state Markov-dependent process (2Mp): return periods T_W and T_N as function
 577 of T for several values of the correlation coefficient ρ in absolute value (a) and normalized
 578 with respect to the independent value T (b). Note that $T_N = T$ for every value of ρ , while
 579 $T_W = T_N = T$ for $\rho = 0$ (black line).

580 3. Two state Markov-dependent process (2Mp): distribution functions of the waiting
 581 time, F_W (a) and of the interarrival time, F_N (b) for $p = 0.9$ and for several values of the
 582 correlation coefficient ρ ; the averages of the distributions (return periods) are indicated by
 583 the vertical dashed lines. For the sake of clarity, the distribution functions of the discrete
 584 random variables W and N are represented as continuous functions.

585 4. Return periods T_W and T_N as function of T and for two values of the correlation
 586 coefficient ρ for the AR(1) process (continuous lines) compared to the two state Markov-
 587 dependent process (2Mp, dashed lines). Note that $T_N = T$ for every value of ρ , while
 588 $T_W = T_N = T$ for $\rho = 0$ (black line).

589 5. AR(1) process: mean conditional waiting time $T_{W|t_e}$ (continuous lines) as function
 590 of the elapsing time t_e for $p = 0.9$ and for several values of the correlation coefficient ρ ;
 591 the corresponding mean unconditional waiting times T_W (dashed lines) are depicted as
 592 reference.

593 6. Probabilities of failure $R_W(T_W)$ (continuous lines), $R_{W|t_e}(T_{W|t_e})$ (dot-dashed lines)
 594 or $R_N(T_N)$ (dashed lines) as functions of p for both AR(1) (a, b) and 2Mp (c, d); graphs

595 refer to the cases $\rho = 0.75$ (a, c) and $\rho = 0.99$ (b, d). Note that $R_{W|t_e}(T_{W|t_e})$ of 2Mp is
596 valid for any $t_e > 0$. Results are compared to the independent case (black line).

597 7. Equivalent Return Period (*ERP*), based on the interarrival time N , as function of
598 the independent return period T for several values of the lag-1 correlation coefficient ρ ;
599 curves for $\rho < 0.75$ are not shown because the differences between *ERP* and T are small
600 to negligible.













