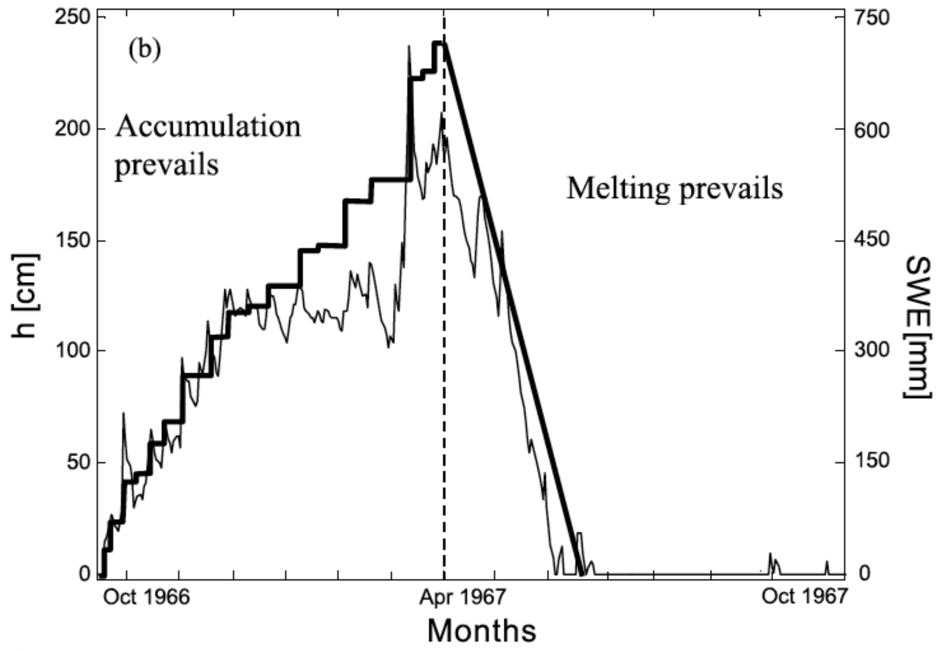
Water Resources Engineering and Management

(CIVIL-466, A.Y. 2024-2025)

5 ETCS, Master course



Lecture 7-3: Stochastic modelling of rainfalls and flows (stochastic models in continuous time)

Prof. P. Perona

Platform of hydraulic constructions

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

EPFL

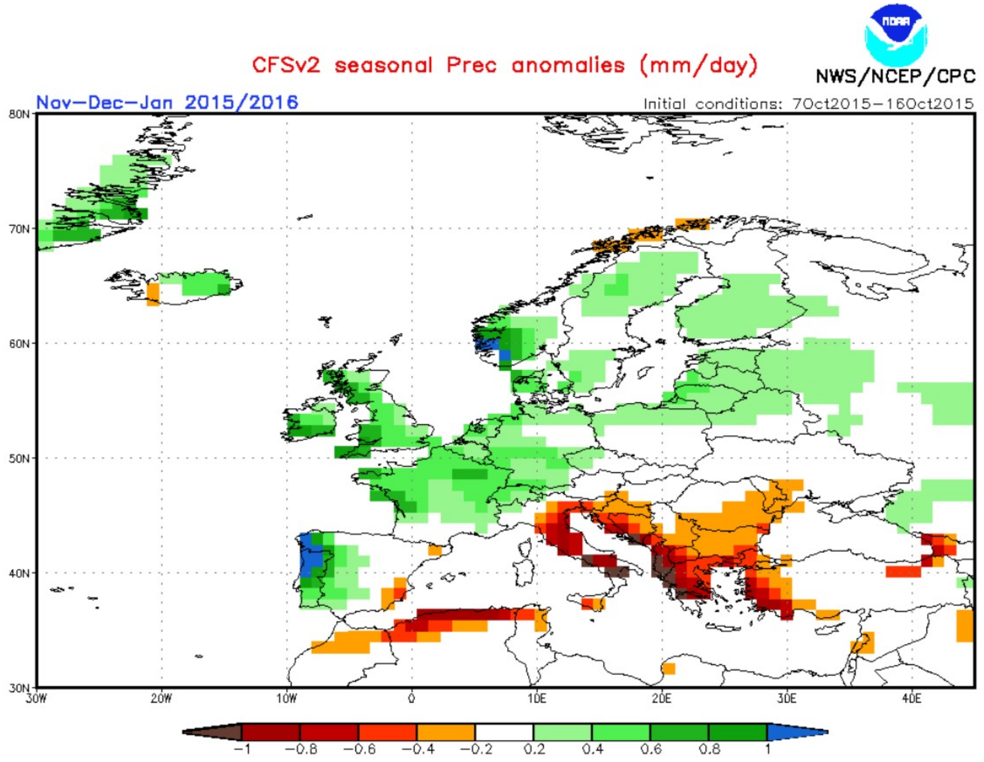
Precipitation modelling at daily (single pulse) and hourly (rectangular pulse) time scales (disaggregation)

Type of rainfall models

Rainfall is the main input of fresh water resources systems. In many water resources engineering and management applications, rainfall models are used to forecast or estimate the amount of resources available or the magnitude of possible floods.

→ Two main classes of rainfall models: deterministic and stochastic models

Deterministic, physically based models: solution of the thermodynamic and fluid-mechanic laws of the atmosphere to predict the future evolution of the state of the atmosphere (Temperature and humidity) given a set of initial conditions (e.g. global circulation models, regional circulation models). Outlook from few days for weather forecast to some months for seasonal forecast of climatic trends (anomalies). In water resources engineering they are used to optimize the short term flood control practice (weather forecast) or to predict the resources available for the next season (seasonal climate forecast).

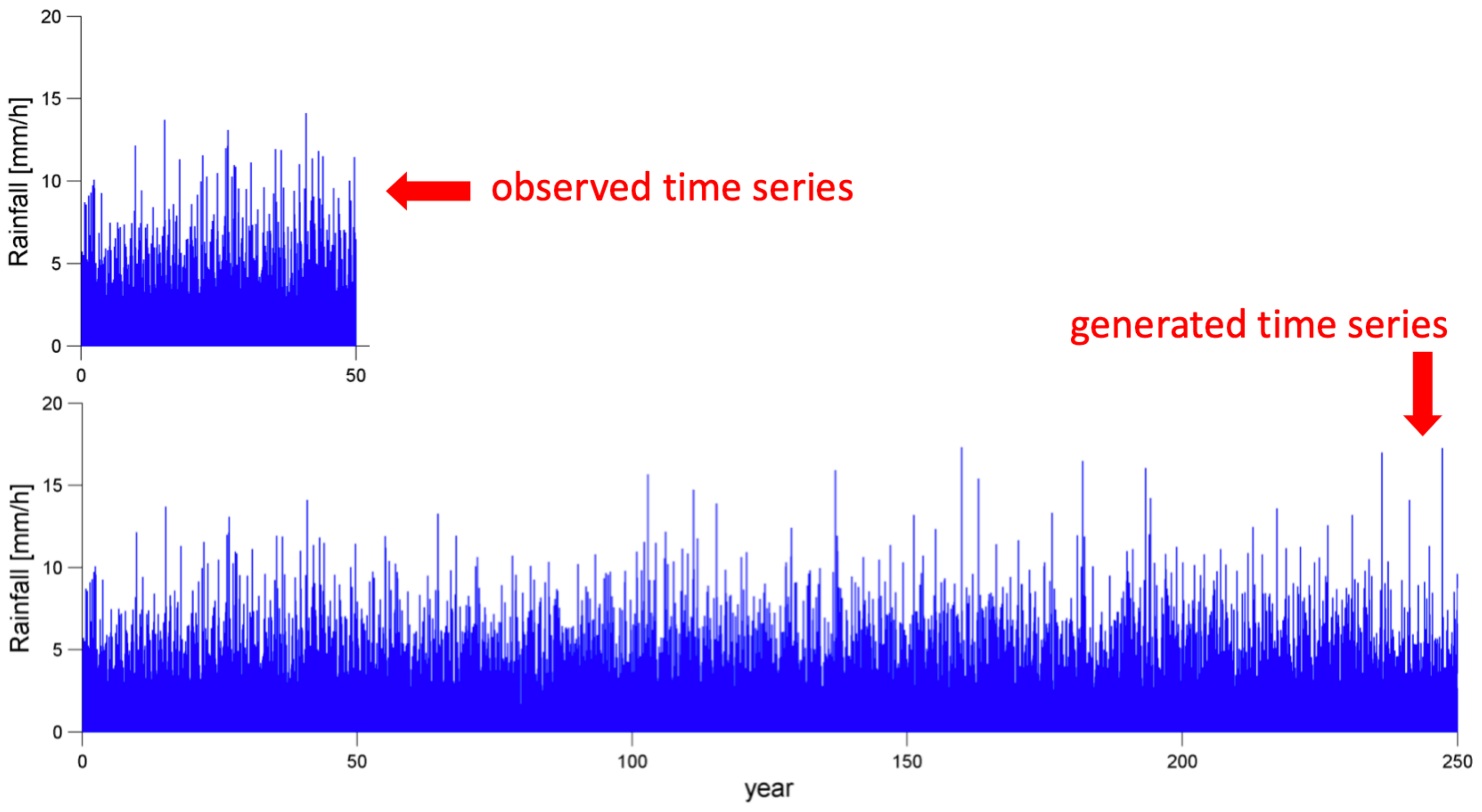


NOAA climate prediction center.
Seasonal precipitation anomalies
Anomalies: variation with respect to the long term average.

Stochastic models: the physical processes are conceptualized (phenomenological approach), the randomness of the underlying phenomena is explicitly included via the definition of random variables having a clear physical meaning. Forecasting is allowed in a statistical sense (statistically meaningful synthetic time series can be obtained). Suitable for long term analysis of water resources systems in a Monte Carlo framework (e.g. estimate the probability of a certain flood, probability of drought and corresponding irrigation water withdrawal, hydropower production, etc).

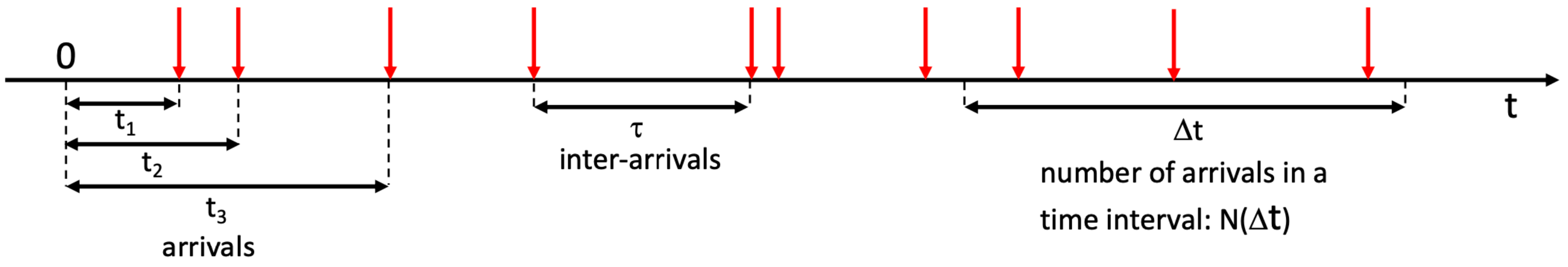
This class of models will be the focus of this lecture.

They can be subdivided into point models (no spatial description of rainfall fields) and spatially explicit models.



Some statistical properties of the Poisson process

Poisson process: stochastic model of random occurrence of instantaneous events: e.g. rainfall events, arrivals of customers in a store, page requests in web server, genetic mutation, etc. etc.



4 equivalent definitions

λ:rate of the Poisson process $[T^{-1}]$

the probability P of having an event in an infinitesimal timespan dt is

$$P = \lambda\, dt$$

inter-arrivals are exponentially distributed

$$f(\tau) = \lambda e^{-\lambda \tau}$$

k-th arrival is gamma distributed

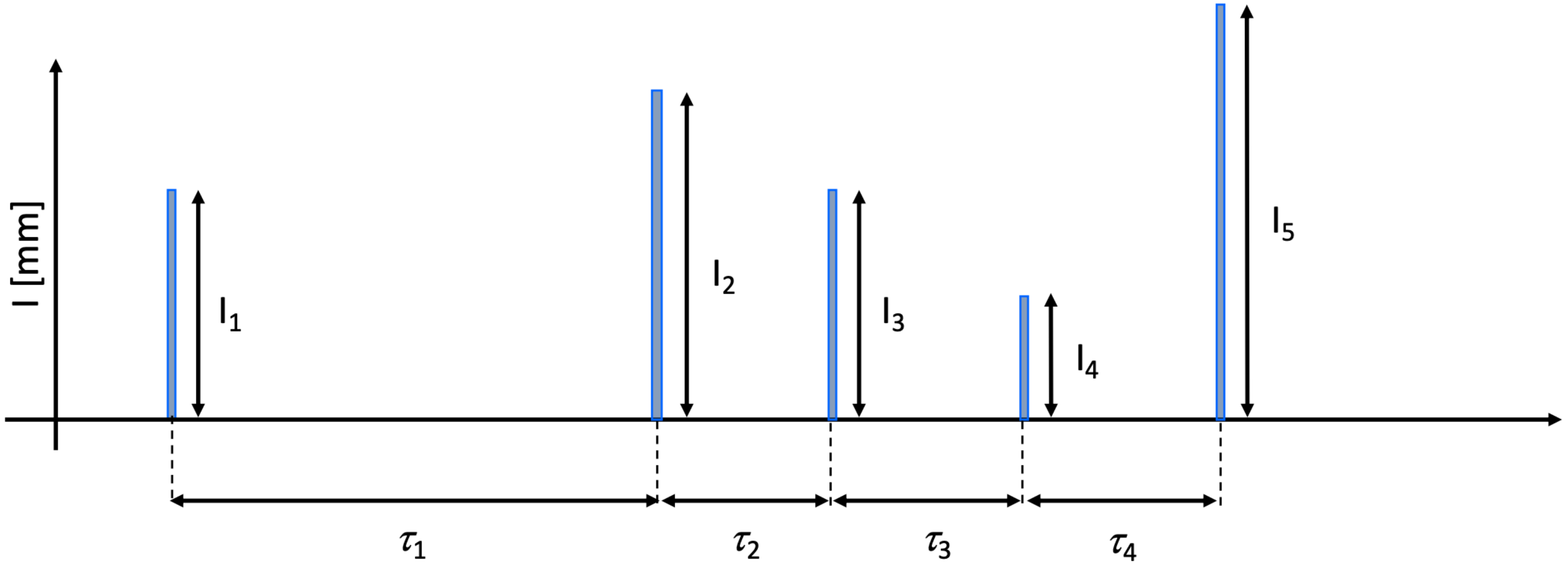
$$f(t_k) = \frac{(\lambda t)^{k-1}}{(k-1)!}\lambda e^{-\lambda t_k}$$

number of arrivals in the time interval Δt follows a Poisson distribution

$$P(N(\Delta t) = n) = \frac{\lambda \Delta t^n}{n!} e^{-\lambda \Delta t}$$

All these statistics are analytically exact

Single pulse rainfall process (Marked Poisson process)



Point model of rainfall. At daily timescale rainfall events are considered uncorrelated (no autocorrelation). The model neglets the temporal variability of rainfall within an event. Rainfall events are modelled as a Poisson process.

The process is described by 2 random variables (exponentially distributed)

τ: event inter-arrival
I : precipitation

$$f(\tau) = \lambda e^{-\lambda\tau} \longrightarrow \langle \tau \rangle = 1/\lambda$$ mean event inter-arrival, λ: Poisson rate

$$f(i) = \frac{1}{\alpha} e^{-\frac{i}{\alpha}} \longrightarrow \langle I \rangle = \alpha$$ mean precipitation

This model is typically used to generate rainfall at a daily time scale.

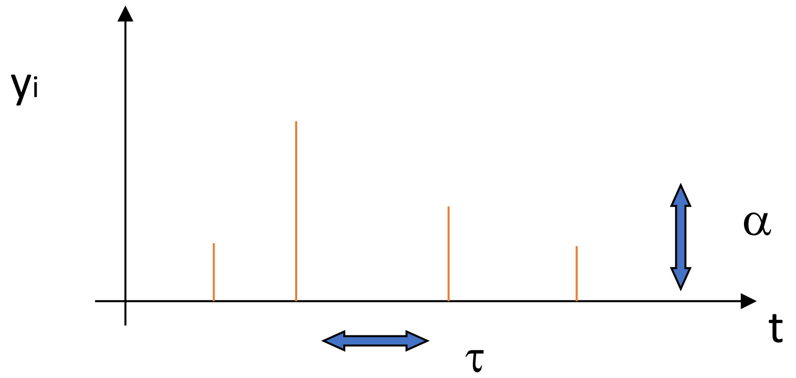
Marked Poisson process (additional remarks)

For a **Poisson process**, which consists of independent increments and follows a purely random (memoryless) process, the Hurst exponent is H=0.5

Explanation:

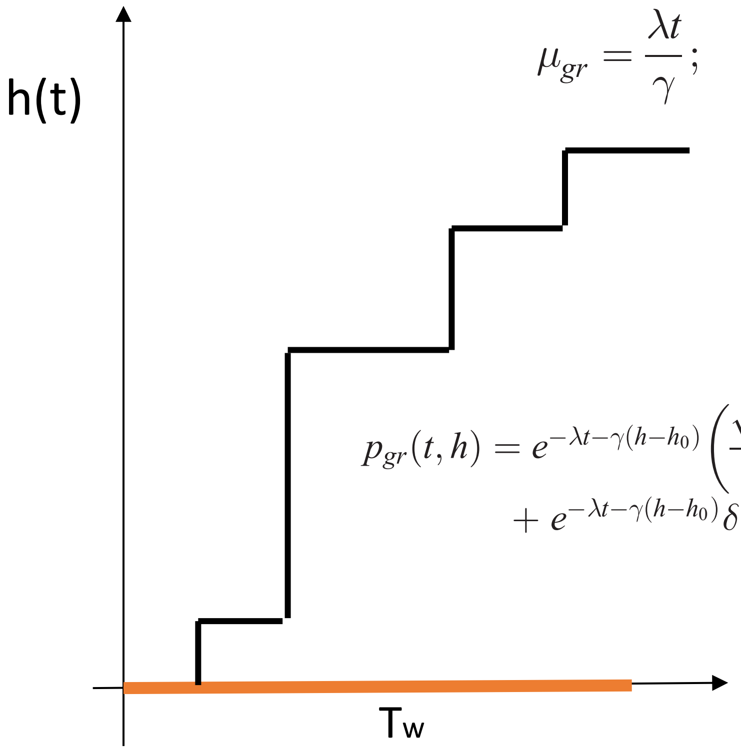
- The Poisson process N(t) has **independent increments**, meaning past values do not influence future values.
- The variance of its increments grows linearly with time, similar to a Brownian motion (which also has H=0.5).
- Since H=0.5 corresponds to a **purely random** process without long-range dependence, the Poisson process falls into this category.

Thus, the **Hurst exponent of a Poisson process is 0.5**, indicating a lack of long-term memory and purely stochastic behavior.



The cumulative process is also known

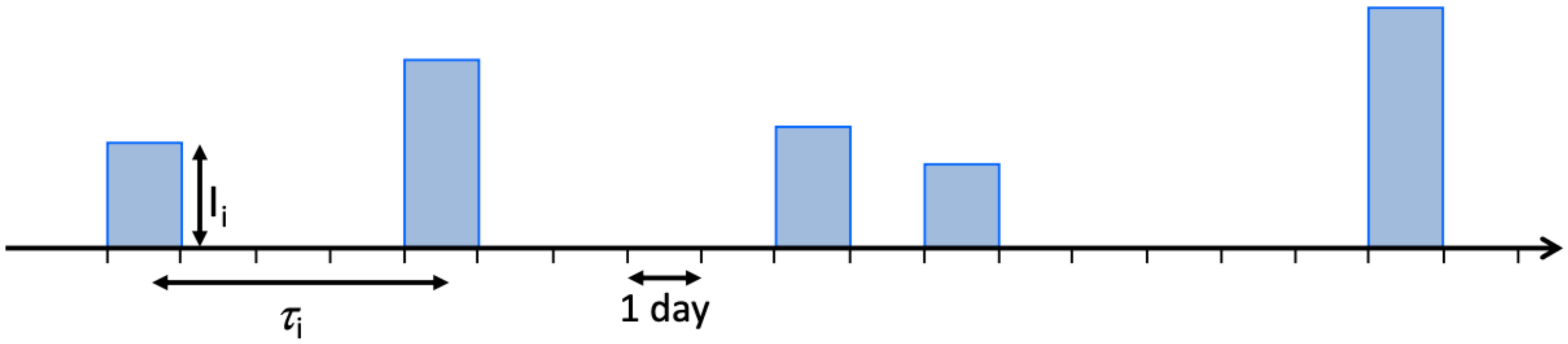
$$\mu_{gr} = \frac{\lambda t}{\gamma}; \qquad \sigma^2_{gr} = \frac{\lambda t(2+\lambda t)}{\gamma^2}.$$



$$p_{gr}(t,h) = e^{-\lambda t - \gamma(h-h_0)}\left(\frac{\sqrt{\gamma\lambda t}I_1\left(2\sqrt{\gamma\lambda t}\sqrt{h-h_0}\right)}{\sqrt{h-h_0}}\right) + e^{-\lambda t-\gamma(h-h_0)}\delta(h-h_0), \quad h \geq h_0$$

Calibration of the single pulse rainfall model

calibration: analysis of the observed rainfall to calibrate the model parameters



direct estimate of the two parameters (mean inter-arrival and mean precipitation)

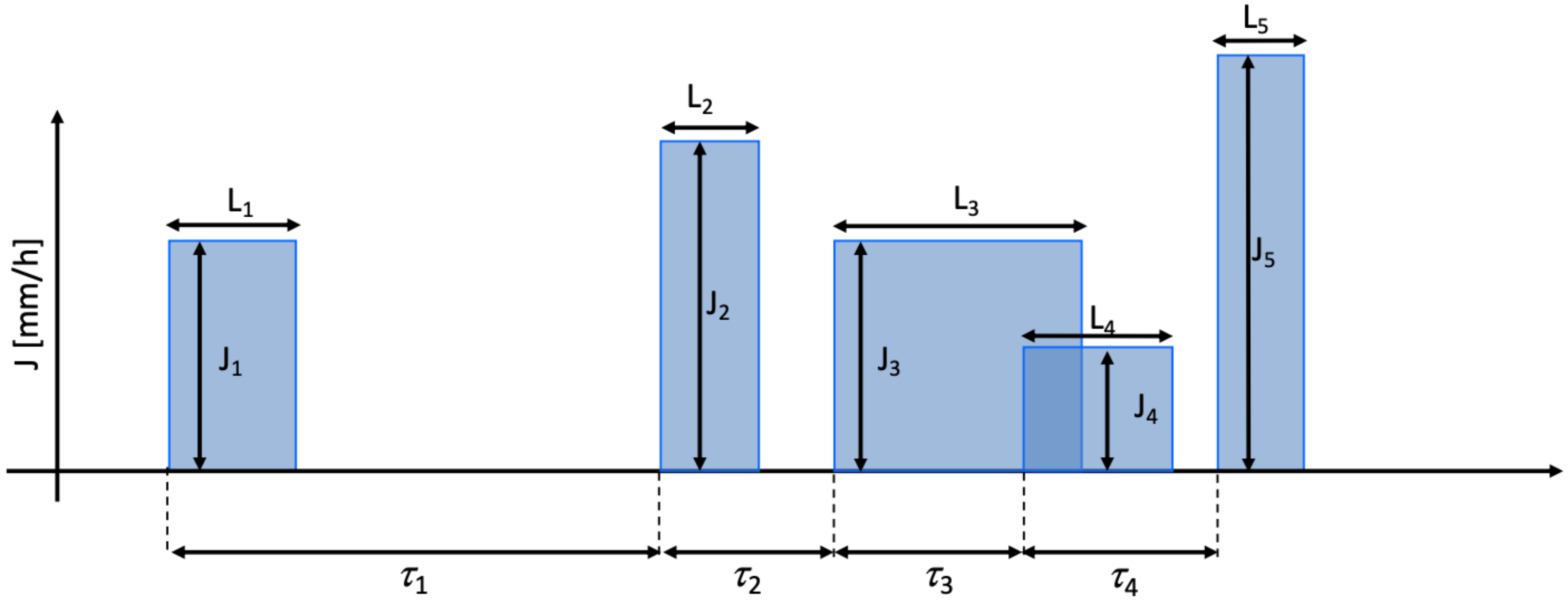
$$\langle \tau \rangle = \frac{\sum \tau_i}{n.of\ rainy\ days} = \frac{n.of\ days}{n.of\ rainy\ days} \longrightarrow \lambda = \frac{1}{\langle \tau \rangle} = \frac{n.of\ rainy\ days}{n.of\ days}$$

$$\alpha = \frac{\sum I_i}{n.of\ rainy\ days}$$

To account for the intra-annual variability of rainfall (i.e. seasonality), the year is divided into periods and different parameters α and λ are estimated for each period. Typically, one month periods are used.

Rectangular pulse rainfall model

Rationale. Precipitation amount at small time scales depends also on the duration of single events (e.g., IDF curves…). The single pulse model is therefore not sufficient anymore

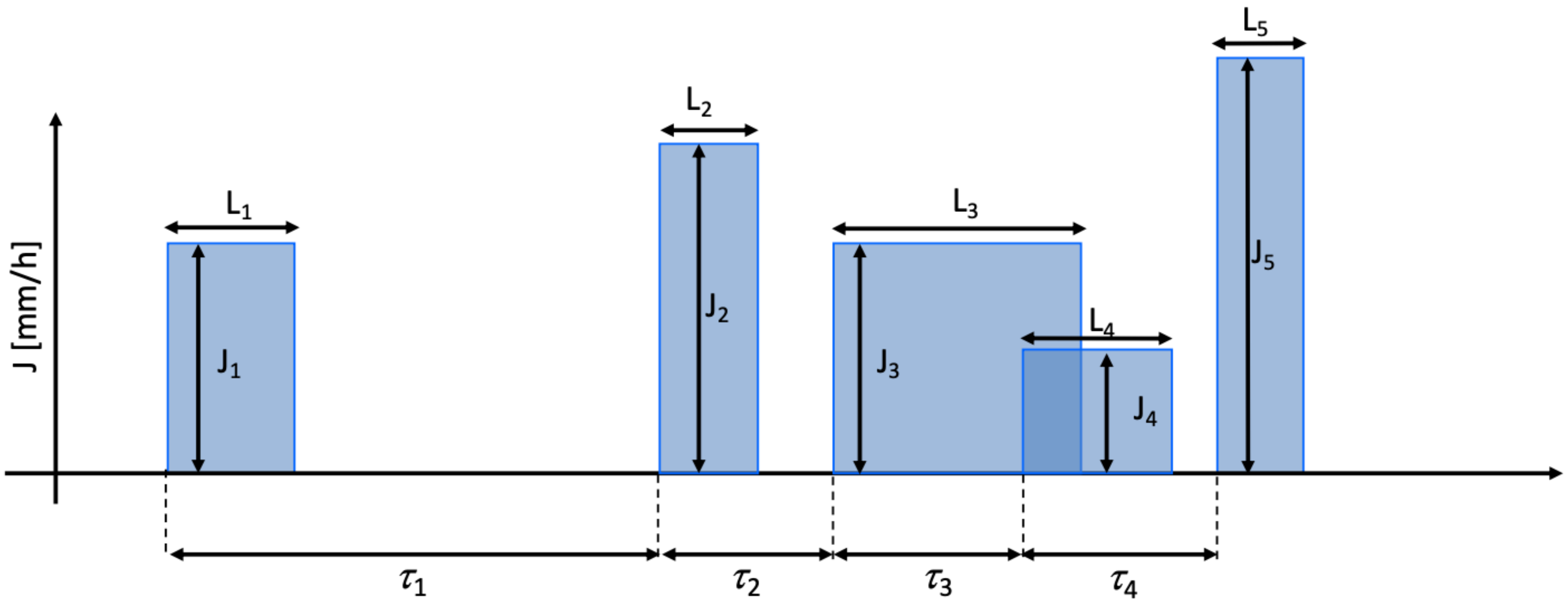


Point model of rainfall. Continuous time model. Occurence of rainfall cells is modelled as a Poisson process. Each cell is characterized by a duration and a precipitation intensity. The process is described by 3 random variables

τ: event inter-arrival
J: precipitation intensity
L: event duration

This model is typically used to generate rainfall at hourly timescale (crucially important for small basins and urban hydrology)

Rectangular pulse rainfall model



All variables are now assumed to be exponentially distributed (3 parameters) and uncorrelated

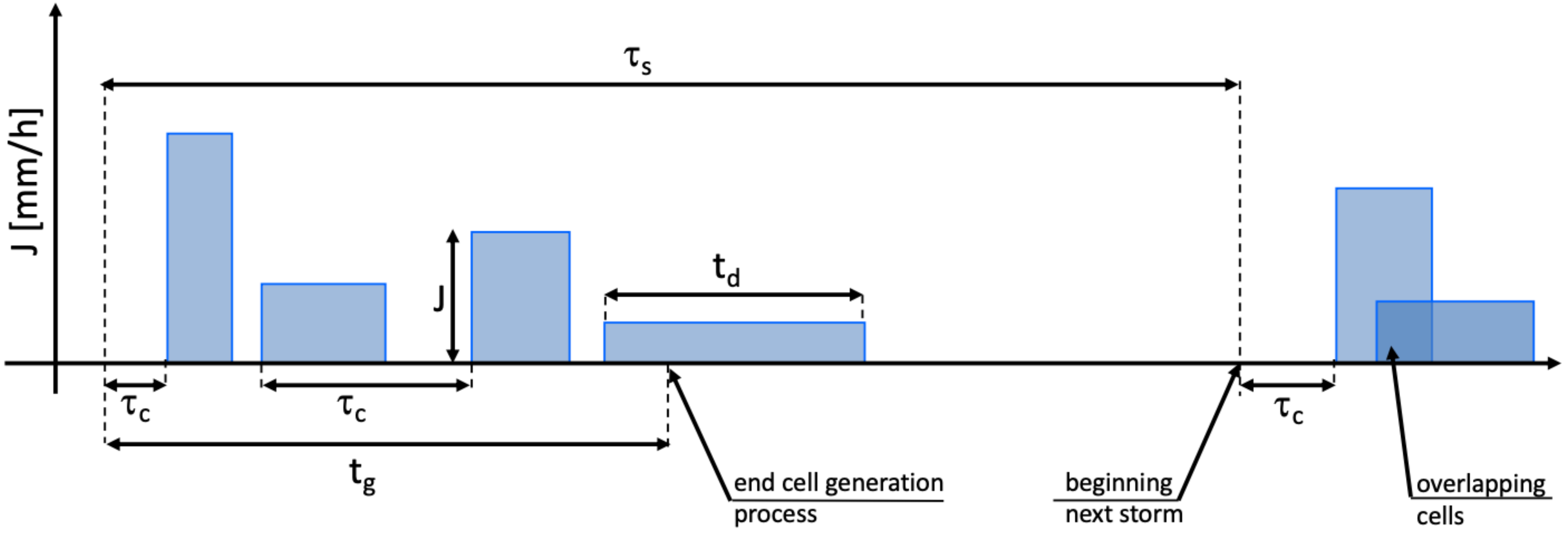
$$f(\tau) = \lambda e^{-\lambda\tau} \longrightarrow \langle\tau\rangle = 1/\lambda$$ mean event inter-arrival, λ: Poisson rate

$$f(j) = \gamma e^{-\gamma j} \longrightarrow \langle j\rangle = 1/\gamma$$ mean precipitation intensity

$$f(l) = \omega e^{-\omega l} \longrightarrow \langle l\rangle = 1/\omega$$ mean cell duration

The added degree of freedom given by the cell duration allows to reproduce also the auto-correlation of the observed rainfall (events can indeed cumulate). Issue: this model does not allow to explicitly group per storm events

Bartlett-Lewis pulse rainfall model



Point model of rainfall. Continuous time model. Rainfall occurs in storms characterized by a certain duration. Whitin each storm, rainfall cells are generated as in the rectangular pulse model. The process is described by 5 random variables generically assumed exponentially distributed

τ_s: storm inter-arrival

t_g: generation of cell time

τ_c: cell inter-arrival

t_d: cell duration

J: cell precipitation intensity

5 parameters are the mean values

$$f(\tau_s) = \lambda e^{-\lambda \tau_s} \longrightarrow \langle \tau_s \rangle = 1/\lambda$$

$$f(t_g) = \gamma e^{-\gamma t_g} \longrightarrow \langle t_g \rangle = 1/\gamma$$

$$f(\tau_c) = \beta e^{-\beta \tau_c} \longrightarrow \langle \tau_c \rangle = 1/\beta$$

$$f(t_d) = \eta e^{-\eta t_d} \longrightarrow \langle t_d \rangle = 1/\eta$$

$$f(j) = (1/\mu_x) e^{-j/\mu_x} \longrightarrow \langle j \rangle = \mu_x$$

Bartlett-Lewis model (6 parameters)

The modified Bartlett-Lewis model (6 parameters) has an additional degree of freedom (one more parameter). It assumes that the parameter h characteristic of the cell duration is also a random variable which follows a gamma distribution (two parameters probability density function). This still allows to obtain all statistics of the process in an exact way (not shown) and so to calibrate on observations (6 conditions)

$$E[Y^{24}] = E[Y^{24}]^*$$
$$\mathrm{var}[Y^{24}] = \mathrm{var}[Y^{24}]^*$$
$$\mathrm{cov}[Y_i^{24}, Y_{i+1}^{24}] = \mathrm{cov}[Y_i^{24}, Y_{i+1}^{24}]^*$$
$$p(24) = p(24)^*$$

information at daily time step

$$\mathrm{var}[Y^{48}] = \mathrm{var}[Y^{48}]^*$$
$$p(48) = p(48)^*$$

information at time step equal to 2 days

* observed quantities

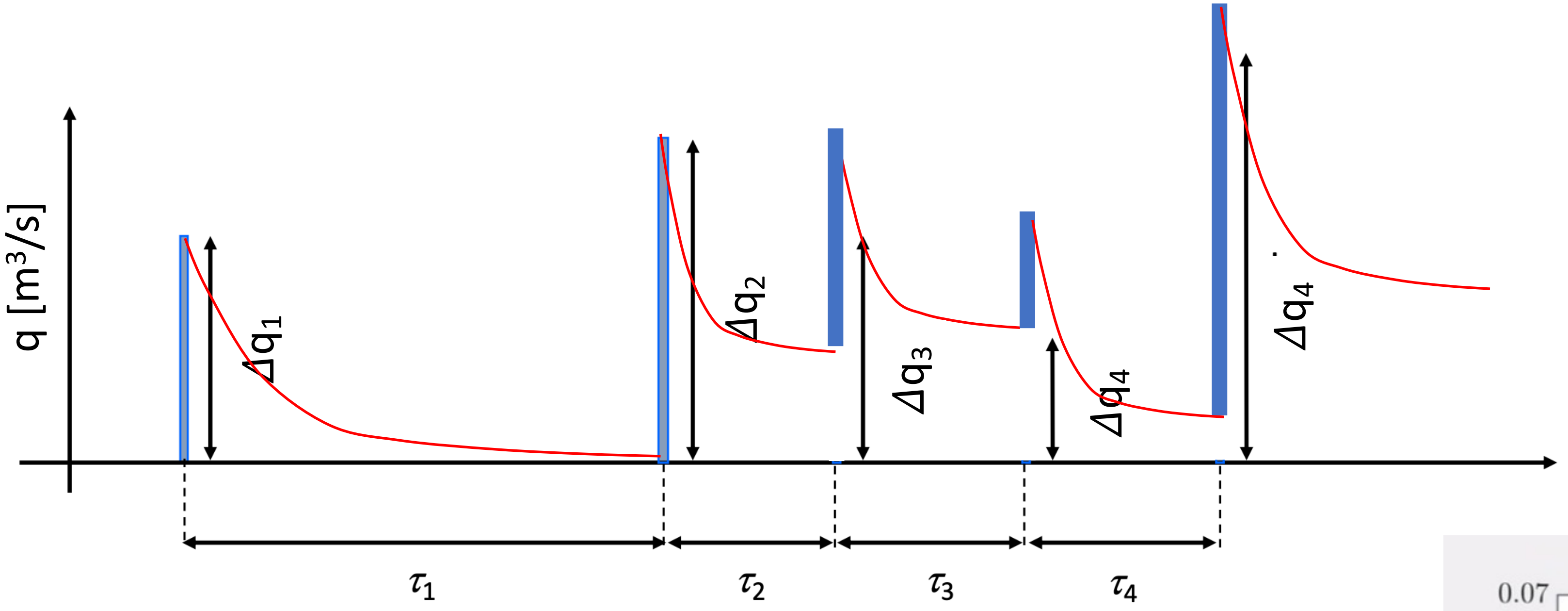
The model reproduces 1) the mean daily rainfall, 2) the daily rainfall standard deviation and 3) covariance, 4) the fraction of dry days; 5) the variance and 6) the fraction of dry periods for a 2- day aggregation.

To reproduce the seasonal variability within the year, data are usually divided into 12 periods corresponding to the months. The calibration is then done for each months, thus obtaining 12*6 parameters!

Rainfall disaggregation is still a very hot topic today and many other models have been proposed (e.g., multifractal, etc.)

Flow discharge modelling and mean hydrograph below and above threshold

Compound Poisson Process

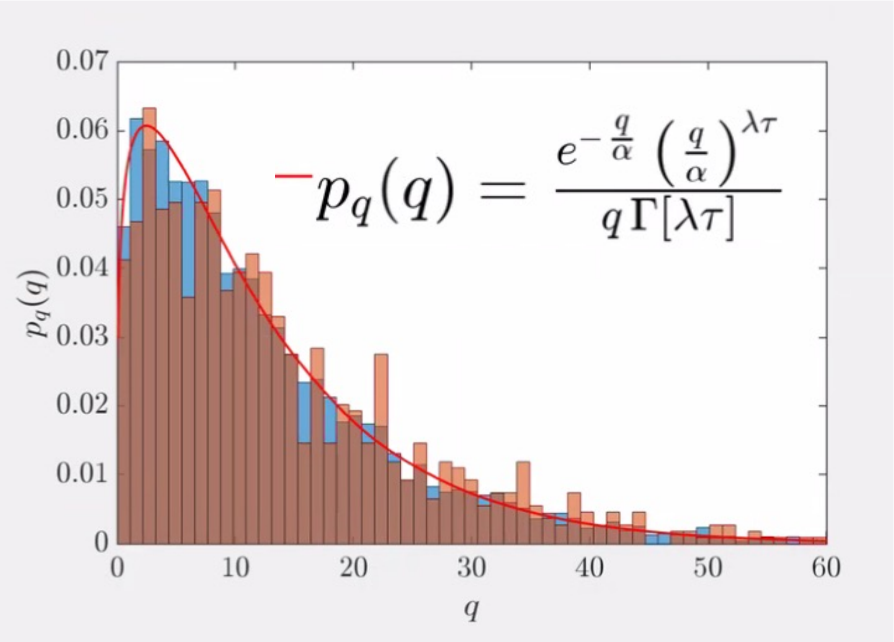


Cumulate the events and decays exponentially between them with a rate $\frac{dq}{dt} = -\frac{1}{\tau} q$,

where $\frac{1}{\tau} = k$ is the decay constant

$f(\tau) = \lambda e^{-\lambda\tau} \longrightarrow \langle\tau\rangle = 1/\lambda$ mean event inter-arrival, λ: Poisson rate

$f(\Delta q) = \frac{1}{\alpha} e^{-\frac{\Delta q}{\alpha}} \longrightarrow \langle\Delta q\rangle = \frac{1}{\alpha}$ mean jump Δq



The resulting signal is now correlated, i.e. a "coloured noise"

This process is used to model river discharges or stages at the daily time scale without trend or seasonality

Take home messages from these three lectures

- L7.1 I remember the structure of (linear) autoregressive model of generic order p and the meaning of the Yule-Walker equation
- L7.1 I can write the structure and the statistics of the AR(1) model and the autocorrelation function (ACF)
- L7.1 I know how to use an AR(1) model to generate new data (generating equation)
- L7.1 I remember the structure of (linear) moving average models and its statistics and the ACF
- L7.1 I remember the structure of the ARMA(p,q) model, but do not need to remember its statistics

- L7.2 I understand how to use the ACF and the PACF for identifying the order of the model
- L7.2 I know how to perform model testing and diagnostic for the AR(1) and MA(1) models
- L7.2 I can explain the structure of periodic autoregressive models but do not need to remember them
- L7.2 I can explain the Thomas-Fiering model (but no need to remember the reqression equations for the coeffs)

- L7.3 I understood the Poisson process and the resulting distributions for the intertime statistics
- L7.3 I can explain the Marked Poisson process and how to generate daily rainfall data
- L7.3 I understand the concept behind the rectangular pulse model and the Bartlett-Lewis model
- L7.3 I can explain the Compound Poisson Process for generating flow discharge data at the daily time scale.