Journal of Hydrology xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Assessing the high frequency quality of long rainfall series

Q1 C.T. Hoang [a], I. Tchiguirinskaia [a,b,*], D. Schertzer [a,c], P. Arnaud [b], J. Lavabre [b], S. Lovejoy [d]

[a] LEESU, University Paris-Est, Ecole des Ponts ParisTech, 6–8 Avenue Blaise Pascal, 77455 Marne la Vallée Cedex 2, France
[b] Cemagref, HOAX, Aix-en-Provence, France
[c] Meteo France, CNRM, Paris, France
[d] McGill U., Physics Dept., Montreal, PQ, Canada

A R T I C L E   I N F O

Article history:
Received 31 December 2010
Received in revised form 23 December 2011
Accepted 31 January 2012
Available online xxxx

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Ercan Kahya, Associate Editor

Keywords:
Long time series
High resolution data
Frequency quality
Scaling analysis
Operational hydrology

S U M M A R Y

High resolution, long and reliable rainfall time series are extremely important to assess reliable statistics, e.g. the Depth–Duration–Frequency curves that have been widely used to define design rainfalls and rainfall drainage network dimensioning. The potential consequences of changes in measuring and recording techniques have been somewhat discussed in the literature with respect to a possible corresponding introduction of artificial inhomogeneities in time series. In this paper, we show how to detect another artificiality: most of the rainfall time series have a lower recording frequency than that is assumed, furthermore the effective high-frequency limit often depends on the recording year due to algorithm changes. This question is particularly important for operational hydrology, because we show that an error on the effective recording high frequency introduces biases in the corresponding statistics. We developed a simple automatic procedure to assess this frequency period by period and station by station on a large database. The scaling analysis of these time series also shows the influence of high frequency limitations on the scaling behaviour, leading to possible misinterpretation of the significance of characteristic scales and scale-dependent hydrological quantities.

1. Introduction

Nowadays hydrology requires reliable statistics for shorter and shorter durations and for a larger and larger range of return periods, which can only be obtained from higher resolution and longer time series (Berndtsson and Niemczynowicz, 1988; Niemczynowicz, 1999; Ogden et al., 2000). For instance, the requirements about temporal and spatial resolutions of rainfall data for urban hydrology have been discussed by Schilling (1991) and Berne et al. (2004) and tentatively quantified. Preliminary studies (Berggren, 2007; Olofsson, 2007) showed that the estimated number of floods was lower when using low time resolution data of high intensity rainfall events, compared to those obtained with the help of higher time resolution data.

However, the recording methods, which are very often based on the idea of a sequence of so-called "homogeneous" rainfall episodes, i.e. episodes with a more or less constant rainrate, and the effective measurement frequency may introduce quality problems in these series (Fankhauser, 1997, 1998; LaBarbera et al., 2002), and therefore spurious estimates, e.g. spurious scaling breaks, that may have drastic consequences for operational hydrology that is more and more focused on short durations estimates. For instance, until 1985 in France and in many other European countries, most of the precipitation data were recorded on paper charts. As shown by (Kvicera et al., 2005) the same paper chart can be deciphered by experts in a significantly different manner. Indeed, the general method of chart reading corresponds to a transcription of the record graph into a series of segments of nearly "constant" slopes, which would correspond to a series of so-called "homogeneous rainfall episodes", which are supposed to have a constant rainrate. However, criteria defining a slope as being constant belong to the domain of pure "human" expertise, therefore this decomposition into a series of homogeneous episodes is always questionable. Furthermore, a precise rainfall measurement during extreme events remains a complicated task due to numerous jumps on the chart. The transition to electronic recording intended to significantly improve the precision of high frequency data. Unfortunately, the compressed data storage and corresponding data pre-processing have remained rather the same and therefore have maintained uncertainties similar to that of the chart transcription. The presently available data correspond to a compression of the original tipping bucket series, i.e. Meteo-France regularly transforms the original data into a series of episodes with a rain rate that is considered as constant in a bracket of 10% that abusively called

* Corresponding author at: Laboratoire Eau Environnement Systemes Urbains, University Paris-Est, Ecole des Ponts ParisTech, 6–8, Avenue Blaise Pascal, Cité Descartes, 77455 Marne la Vallée Cedex 2, France. Tel.: +33 01 64 15 36 23; fax: +33 01 64 15 37 64.
E-mail addresses: Tuan.Hoang@enpc.fr (C.T. Hoang), Ioulia.Tchiguirinskaia@enpc.fr (I. Tchiguirinskaia).

0022-1694/$ - see front matter © 2012 Published by Elsevier B.V.
doi:10.1016/j.jhydrol.2012.01.044

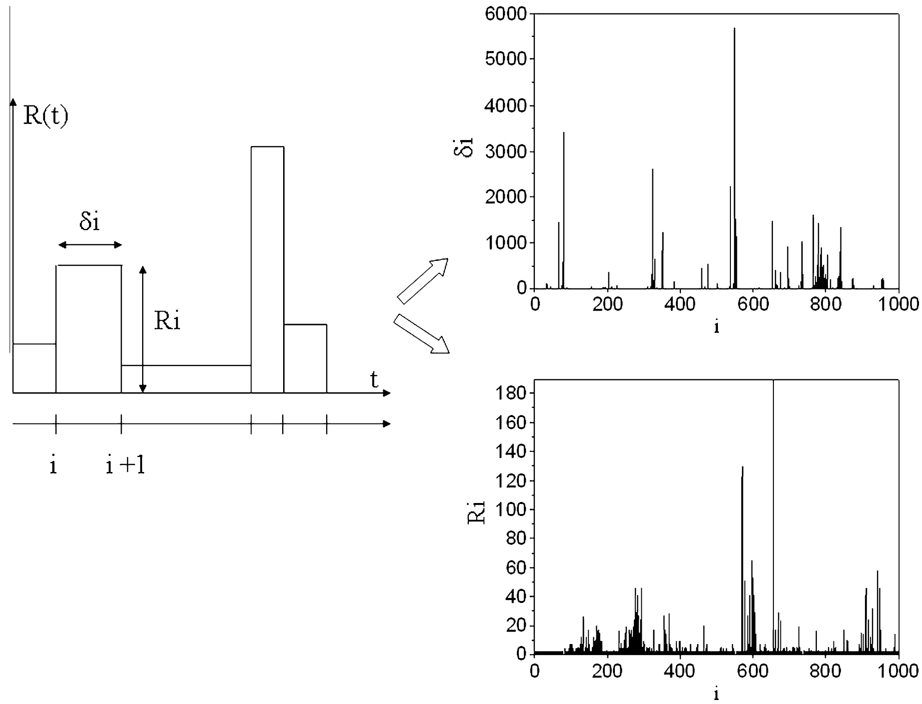

Fig. 1. Illustration of the extraction of the « homogeneous » rainfall episodes from the Rimbaud original time serie.

“homogeneous” rainfall episodes. As illustrated on Fig. 1, the original time series of rainfall are first transformed into a series of successive episodes. The duration of the homogeneous episodes are multiple of respectively 5 and 6 min for MF-P5 and MF-P6. An obvious advantage of such a data compression is a significantly reduced data volume compared to rainfall series stored with a constant time increment. The series of homogeneous episodes therefore corresponds to a coupled series of discrete rainfall durations δ_i and discrete rainfall intensities R_i ($i = 1,N$). Both series display a strong variability that, to some extent, has an opposite meaning (see Fig. 1): highest values of duration δ_i generally correspond to the “zero” rainfall, i.e., that is under a level of detection by the actual tipping bucket device; whereas the highest rainfall intensities R_i are observable over a rather short durations δ_i.

A preliminary study of a set of 10 time series from a French database (Hoang, 2008) put in evidence the deficit of short duration episodes. It seems that the understanding to some extent of this deficit resulted in some cases in a partial reconstruction of higher frequency rainfall time series, while such data were still available. The time series having mixed data frequencies of measurements may explain often observed scaling breaks that we will discuss below.

Because the importance of scales and possible scaling behaviour of hydrological data is particularly important for many applications (Tchiguirinskaia et al., 2004; Aronica and Freni, 2005; Kundu and Bell, 2006), this paper investigates the sensitivity of the scaling estimates and methods to the deficit of short duration rainfall data, and consequently propose a few simple criteria for a reliable evaluation of the data quality.

2. Data case study

Our study bears on 166 long rainfall time series that are a part of a Meteo-France database (MF-P5) that have been used to calibrate the hourly hydrological SHYPRE model (Arnaud and Lavabre, 1999) and to estimate the long return period quantiles at hourly and longer durations. A particularity of the MF-P5 database is that 5 min rainfall accumulation estimates being available due to particular measuring experiments over some limited periods of time, were integrated in the originally hourly rainfall database. The more recent database (MF-P6) is based on 6 min rainfall accumulation estimates. In this paper we use three available rainfall time series of MF-P6 (Brest, Mont Aigoual and Marseille) that correspond to the respective periods of measurements from 1990 to 2008, from 1992 to 2008 and from 1982 to 2008. Ongoing research is devoted to statistically estimate and stochastically simulate reliable subhourly rainfall quantiles, in part with the help of the above databases. Thus the reliable evaluation of the data quality at shorter durations is particularly indispensable.

The locations of these 166 gauges are rather evenly distributed over France, although with a higher concentration over regions of particular interest for flood studies, e.g. the Mediterranean area. The length of these time series ranges from 9 to 88 years.

As a preliminary data analysis, we computed the probability of the durations δ_i of non-zero rainfall episodes. As illustrated by Fig. 2, some of the duration probabilities are clearly dominated by only a few (or even a unique!) characteristic durations. The episode duration having the highest occurrence probability corresponds to one of the three following cases.

The first one (e.g. the time series of Nîmes, Fig. 2a) merely corresponds to hourly data sets, that were recorded with the format of homogeneous episodes, without being actually transformed into such episodes. Although, there are rather surprisingly two exceptions: a few rare episodes have durations of 5 min (0.04%) and 115 min (0.02%) respectively.

The second case (e.g. the time series of Marseille, Fig. 2b,) also corresponds to the dominance by the hourly duration (30.2%), but with other durations that are less negligible than in the case of the Nîmes times series, e.g. the Marseille time series has 5.7% of homogeneous episodes with a duration of 5 min. Furthermore, the full histogram of the Marseille series (Fig. 2b) displays durations that are multiples of 5 min, including durations longer than 1 h. One can therefore suspects that the Marseille time series, con-

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

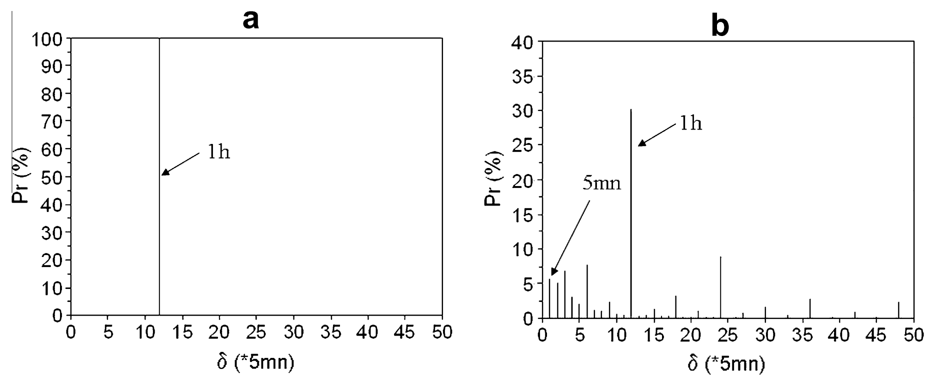

Fig. 2. Probability distribution of the durations of « homogeneous » rainfall episodes in the Nîmes (a) and Marseille (b) time series. Both series have rather hourly effective resolution.

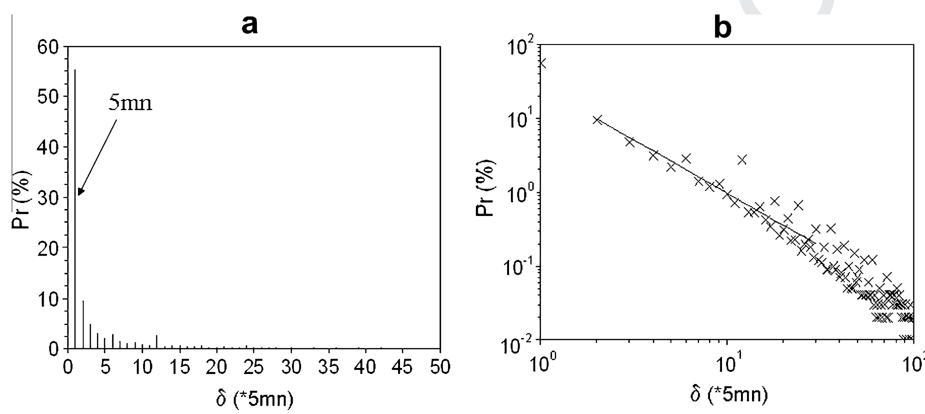

Fig. 3. Probability distribution of the durations of « homogeneous » rainfall episodes in Rimbaud time series in a linear plot (a) and a log–log plot (b). The series has an effective 5 min resolution.

trary to the Nîmes series, were at least partially constructed from 5 min rainfall data with a given algorithmic transformations, which in turn might have artificially increased the duration of homogeneous episodes. Unfortunately, no document seems to be available about this presumed transformation algorithm.

Finally, the third and most straightforward case corresponds to the case when the smallest durations are dominant. The time series of Rimbaud is one of a few rainfall data series that displays (see Fig. 3) a clear peak of the duration probability at 5 min.

What can we infer from these rather different behaviours of the duration probability? From a scaling point of view, a peak of the duration probability at the shortest available duration, i.e. the recording duration, is rather natural, because rainfall has higher and higher variability over smaller and smaller time scales. One may furthermore expect that the relation between the duration of rainfall episodes with a given relative homogeneity and its probability of occurrence should have no characteristic scale and therefore should be a power law. This has been verified on multifractal rainfall simulations. For example, Fig. 4 displays a power law behaviour, emphasised by a linear fit in logarithmic coordinates, of a probability distribution of 10% “homogeneous” episode durations, obtained on a rainfall simulation with multifractal parameters $\alpha = 1.5$ and $C_1 = 0.15$. The universal multifractal parameters C_1 and α (Schertzer and Lovejoy, 1987b, 1997) measure respectively the mean intermittency of the rainfall ($C_1 = 0$ if the rainfall is homogeneous or, loosely speaking, if it always rains) and the variability of the rainfall intermittency ($\alpha = 0$ if the intermittency of the extremes is the same as the intermittency of the mean rainfall). In fact, the power law behaviour of the distribution depends on the

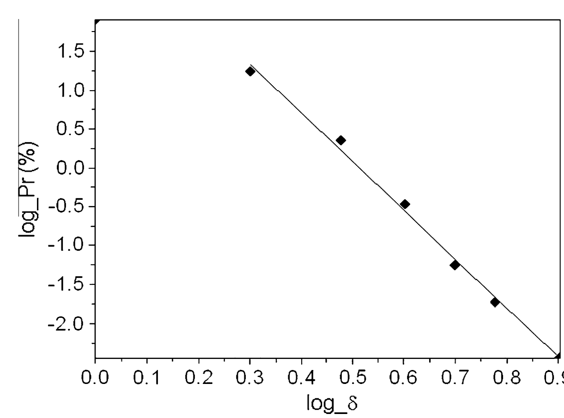

Fig. 4. Log–log plots of a probability distribution of the durations of 10% “homogeneous” episodes of a multifractal rainfall simulation ($\alpha = 1.5$ and $C_1 = 0.15$). It displays a rather clear power law behaviour that corresponds to a good linear fit in a log–log plot.

required degree of homogeneity and on the values of the multifractal parameters. This power law behaviour is better respected with strong homogeneity thresholds, e.g. 5% and 10%. It becomes questionable with weaker homogeneity thresholds, e.g. 20% and 40%. On the other hand, for a fixed α-value, a power law behaviour is more evident for higher values of C_1. Overall, similar power law behaviours are therefore expected for measured rainfall data from the MF-P5 and MF-P6 databases. For instance plotting now the duration probability of the Rimbaud time series (Fig. 3a) in

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

logarithmic coordinates (Fig. 3b) we do obtain a linear behaviour, which is particularly obvious over the smaller time scales.

Overall, these three examples illustrate first that databases are not always homogeneous in the sense that the time series have not a uniform measurement or recording frequency. Therefore, a preliminary data quality analysis is rather indispensable before using this database down to its claimed highest time resolution. This is particularly indispensable for the hydrological estimates being based on the scaling analysis. Indeed, while in the third case scaling regimes could be expected to hold over the full range of scales; the scaling behaviour will be presumably broken in the second case, due to small scale data deficit, and certainly broken in the first case due to the absence of small scale data. Contrariwise, scaling techniques can be useful to assess the data quality.

Therefore, in the next sections we perform a scaling analysis of the MF-P5 database, with a particular emphasis on various estimation problems most presumably related to the data quality. Then we discuss the principles of the SERQUAL procedure, which is designed to routinely assess the quality of time series, helping to select sequences of high quality data and therefore to significantly reduce the dispersion of the hydrological estimates, e.g. quantiles as well as scaling parameters.

3. Scaling analysis

It is rather usual, particularly in turbulence, to use the energy spectrum for a preliminary test of scaling behaviour. Indeed, scaling corresponds to a power-law energy spectrum $S(k)$ (Kolmogorov, 1941; Obukhov, 1941):

$$S(k) \propto k^{-\beta} \tag{1}$$

where k is the frequency and β is the spectral exponent. In logarithmic coordinates, this power-law corresponds to a straight line, whereas one generally obtains a highly spiky spectral behaviour for an individual realisation of the rainfall data. Indeed, for an empirical energy spectrum, the power-law is obtained only for averages over a large number of realisations. When an averaged spectrum still contains significant spectral spikes over given scales, these scales are generally considered as characteristic ones. For example, for time series influenced by the annual cycle, the annual spectral spikes have been often discussed in the literature (Tessier et al., 1993). Spectral spikes over small scales, in particular from around 1–3 h for the rainfall, are rather frequent and feed the ongoing debate on whether there is a small scaling break at these scales (e.g. de Lima, 1998).

Fig. 5 displays outputs of the spectral analysis for the Rimbaud time series. We first performed a spectral analysis over a yearly realisation in order to test the hypothesis of inter-annual seasonality of the rainfall data. Indeed a spectral analysis is a useful tool to detect periodic behaviour in time series: the frequency that corresponds to a periodic time-scale will be characterised by an obvious spectral spike. Indeed, the spectrum of a pure periodic signal with a given frequency is a Dirac function centered at this frequency. As illustrated by Fig. 5a, the spectral exponent of $\beta = 1.087$ approximates well the power law behaviour of the Rimbaud time series. Furthermore, in spite of the large fluctuations of individual spectra around the mean value of spectral exponent, there is not any obvious spectral spike. Therefore, there is no observational evidence of seasonal effects on the monthly ensemble average of spectral estimates. Being interested in scaling behaviour over smaller scales, we then performed the spectral analysis over the monthly sequences of each rainfall time series. The data sequences were used as independent realisations. Hence, for each frequency the estimated average spectral value corresponds to a simple average over all individual realisations. Let us emphasise the fact that we used a common practice of ensemble average with weak dependence between samples: one should not be confused between the data resolution (at which rainfall exhibits a strong dependence) and the much larger sample length (which corresponds to weak dependence for rainfall). Furthermore, the goal of this type of spectral analysis is to uncover the small scale dependence, not the larger scale one. As illustrated by Fig. 5, the interest of such an ensemble average is to easily uncover the spectral exponent by smoothing out the large fluctuations of the individual spectra: the energy spectrum averaged over 12 monthly realisations (Fig. 5b) is much smoother than the spectrum of a unique realisation (Fig. 5c) of the Rimbaud time series. Both are characterised by the approximate spectral exponent of $\beta = 1.08$, although the latter is much more obvious on the averaged spectrum. This also shows that no serious artefact is introduced by the ensemble average, but merely that

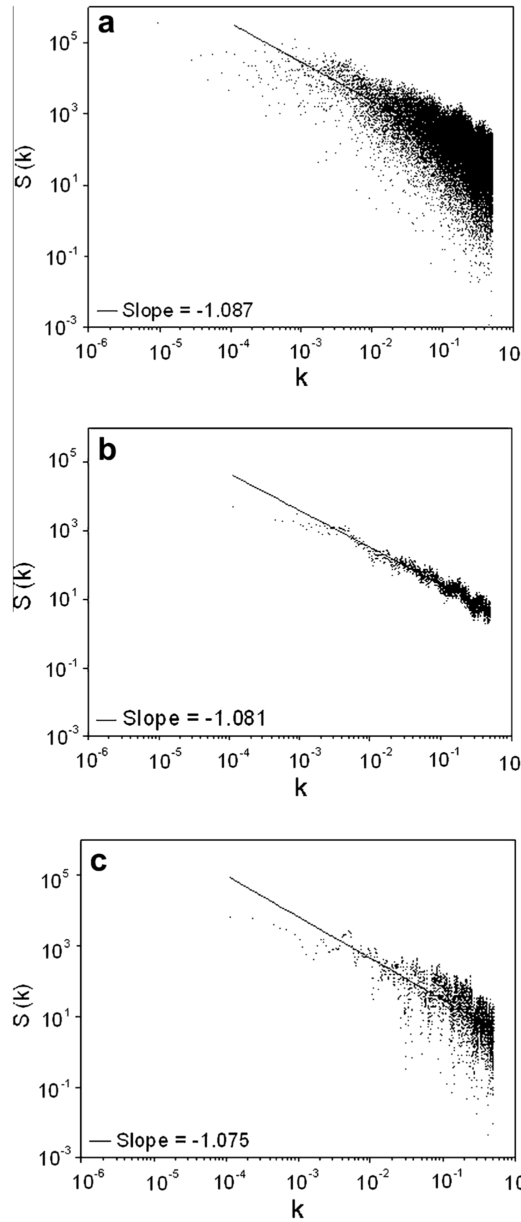

Fig. 5. Log–log plots of the energy spectra of the Rimbaud time series: the spectrum of a yearly data sample (a), the average of the corresponding 12 monthly spectra (b) is obviously smoother than one of the 12 monthly spectra (c). All spectra have nevertheless the same log–log regression that yields the same estimate of the spectral exponent $\beta = 1.08$.

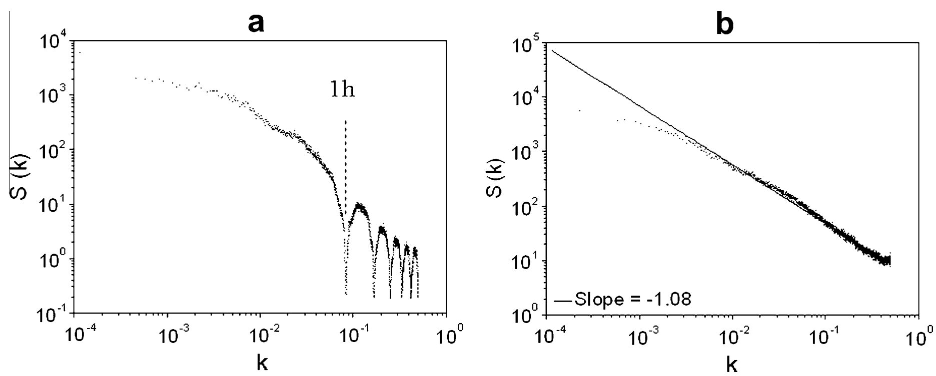

Fig. 6. Log–log plots of the energy spectra of the Nîmes (a) and Rimbaud (b) time series: the latter displays a rather clear scaling behaviour (power law corresponding to a linear fit in a log–log plot), whereas the former does not do it, in particular for frequencies $k \geqslant (1\ \mathrm{h})^{-1}$.

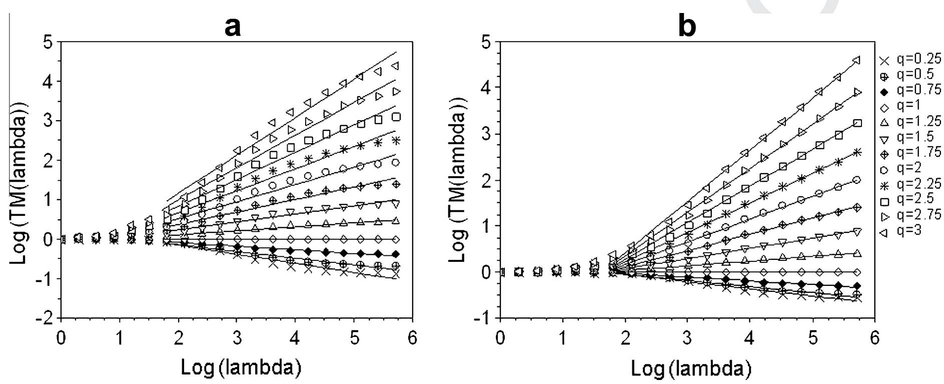

Fig. 7. Log–log plots of the Trace Moments of the Rimbaud (a) and Orgeval (b) time series. Both display a rather clear multifractal or multiscaling behaviour (power laws corresponding to linear fits in a log–log plot), but the scaling holds up to the highest resolution (shortest duration = 5 min.) only for the Orgeval time series.

fluctuations around an average behaviour are smoothed out, as expected. Very close estimates were obtained for all three spectral exponents (Fig 5a–c). This confirms the idea that the monthly ensemble average spectra remain more practical to investigate the scaling behaviour of the data over the small scale range.

Fig. 6 displays the energy spectra of Nîmes and Rimbaud time series. While the Rimbaud rainfall data displays a rather clear scaling behaviour emphasised by a linear fit with slope $\beta = 1.08$ (see Fig. 6b), the small scale data deficit for the Nîmes time series introduces a spurious spectral tail for frequencies $k \geqslant (1\ \mathrm{h})^{-1}$ (Fig. 6a), with apparent scaling breaks at each of the harmonics of the effective recording frequency.

The spectral analysis confirms that the deficit of small scale components, which was first uncovered with the help of the duration probability histograms, indeed introduces high frequency scaling breaks in a spectral analysis of these time series. With this example, it becomes obvious that the hourly data could not be used as a substitute for 5 min rainfall time series. Furthermore, the results suggest that some of the earlier arguments proposed to physically justify the existence of scaling breaks (i.e. breaks in the power laws of the spectrum) at small scales in rainfall may need some re-evaluation from the data quality point of view, to better distinguish real scaling breaks from spurious ones.

Let us note that similar scaling breaks would be observable with the help of different methods of multiscaling analysis of rainfall. Whereas spectral analysis corresponds to a second order statistical analysis, the multifractal Trace Moment (TM) analysis (Schertzer and Lovejoy, 1987a) is performed over a range of orders q of statistical moments. These statistical moments can be estimated with the help of a possible combination of both ensemble and time/space averages, i.e. in agreement with the law of large numbers, theses estimates of the statistical moment of order q correspond to averages of the corresponding power q of the rain rate R_λ over a wide range of resolutions λ (=T/d, where T is the largest time scale of the scaling regime, d is the observation duration). This allows assessing the scaling behaviour of extremes (corresponding to statistical moments of high orders q's, i.e. q's much larger than the unity) as well as for moderate rainfalls (statistical moments of moderate orders, i.e. q's of the order of the unity). As illustrated by Fig. 7, the multiscaling or multifractal behaviour corresponds to a power law behaviour of the corresponding moments ($\langle . \rangle$ denotes the ensemble average):

$$\langle R_\lambda^q \rangle \propto \lambda^{K(q)} \tag{2}$$

here the rainfall rates R_λ at various resolutions $\lambda < \Lambda$ (Λ corresponds to the highest available data resolution) are obtained by a successive upscalings (aggregations) of the original time series R_Λ, the exponent function $K(q)$ describes the scaling of the q-order moments $\langle R_\lambda^q \rangle$ over a given range of resolution λ's. Where it is nonlinear, it corresponds to multiscaling or multifractal behaviour.

For each time series, we first performed the TM analysis over individual sequences of about 5 years of 5 min rainfall (i.e., 2^{19} values), which were considered, as usually done, as almost independent data realisations (see above discussion on ensemble spectra). Then for each of 166 time series, the value of statistical moment at a given resolution λ corresponds to the average estimated at this resolution over all such realizations. Again, the scaling of the empirical trace moments, with respect to the data

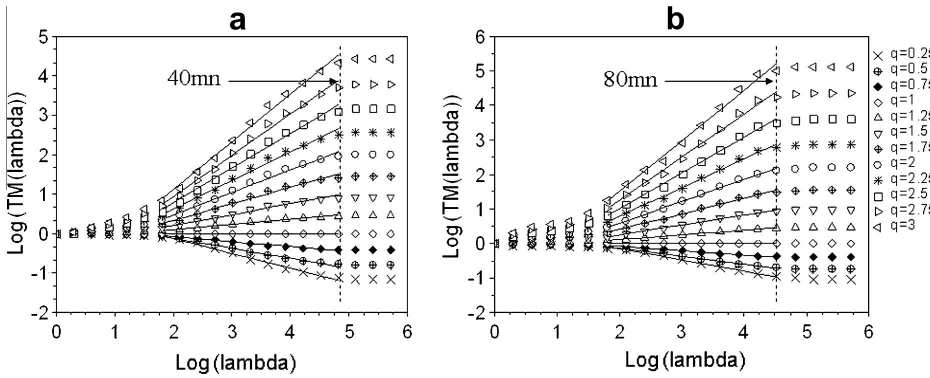

Fig. 8. Log–log plots of the Trace Moments of the time series of Nîmes (a) and Saint Andre de Roquepertuis (b). Both display a clear scaling break at about 1 h and indeed the probability distributions of their durations show hourly effective measurement frequency, instead of 5 min (see Fig. 2a for Nîmes).

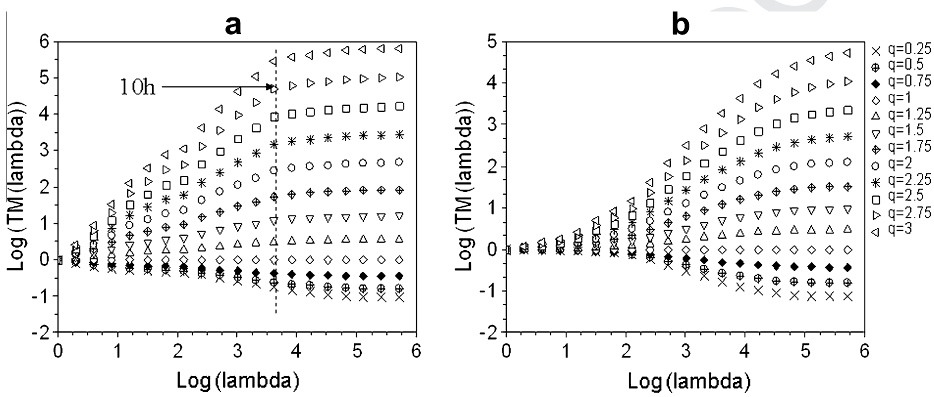

Fig. 9. Log–log plot of the Trace Moments of the Mont Aigoual time series (a) with the scaling break at about 10 h; and Marseille time series (b) without a clear scale of the scaling break, in agreement with Fig. 2b.

resolution λ, corresponds to a linear behaviour of their curves in logarithmic coordinates.

For most of the series, TM curves display scaling breaks (see Fig. 8) at about 40–80 min. As discussed below, they can be considered as a spurious because they could result from the deficit of short "homogeneous rainfall episodes" in the record of series (see Fig. 2 and corresponding discussion). These breaks may occur for larger periods, e.g. at about 10 h for the Mont Aigoual time series (see Fig. 9a), and the period of the scaling break is not clear for the Marseille time series (see Fig. 9b). This obviously illustrates the difficulty and raises many concerns on the possibility to use this time series to benchmark statistical analyses and stochastic simulations.

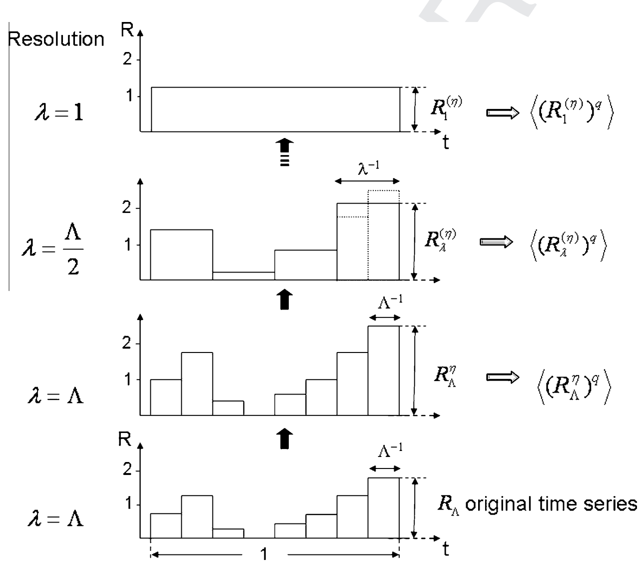

Fig. 10. Schematic illustration of the DTM algorithm that displays its main steps.

A more refined multifractal analysis is obtained with the help of the Double Trace Moment (DTM) analysis (Lavallée et al., 1992; Schmitt et al., 1992). The latter allows to estimate in a rather straightforward manner the 'universal' multifractal parameters C_1 and α that in the case of universal multifractals (Schertzer and Lovejoy, 1987b) fully define the analytical expression of the scaling moment function:

$$K(q) = \frac{C_1}{\alpha - 1}(q^{\alpha} - 1) \tag{3}$$

The main steps of the DTM analysis are schematised on Fig. 10. The first step corresponds to rising up the original time series R_Λ to an arbitrary power η and then proceeding to a TM analysis of the corresponding field R_Λ^η. Therefore, the DTM method estimates the scaling of the statistical moments of $R_\lambda^{(\eta)} (1 < \lambda < \Lambda)$, which are obtained by upscaling (aggregating) R_Λ^η over a scale ratio Λ/λ that generalises the Eq. (2) into:

$$\langle R_\lambda^{(\eta)q} \rangle \approx \lambda^{K(q,\eta)} \tag{4}$$

For every value of the statistical moment order q, Eq. (4) allows to estimate $K(q, \eta)$ as a function of different η-values with the help of a linear regression of $\mathrm{Log}(\langle R_\lambda^{(\eta)q} \rangle)$ vs. $\mathrm{Log}(\lambda)$ over a range of λ 's (see Fig. 11a).

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

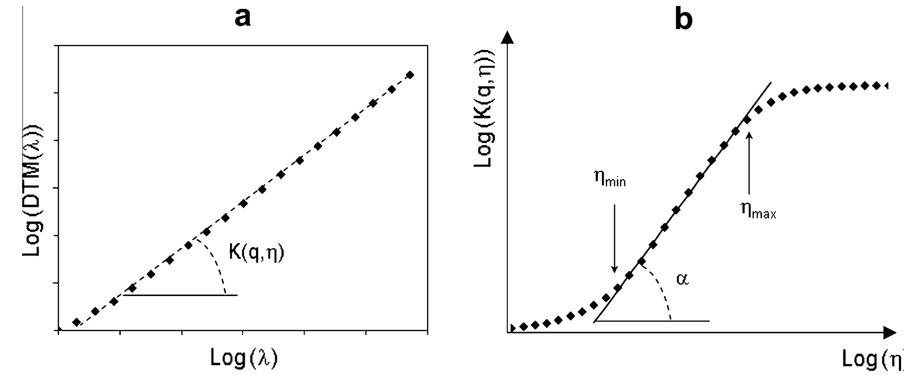

Fig. 11. Illustration of determination of the double trace moment exponent $K(q,\eta)$ values (a) as the slope of $Log(\langle R_{\lambda}^{(\eta)q}\rangle)$ vs. $Log(\lambda)$ and the multifractality index α (b) as the slope of Log($K(q,\eta)$) vs. Log(η) in the DTM analysis, as well as the range $[\eta_{\min}, \eta_{\max}]$ of the η values over which it can be accurately estimated.

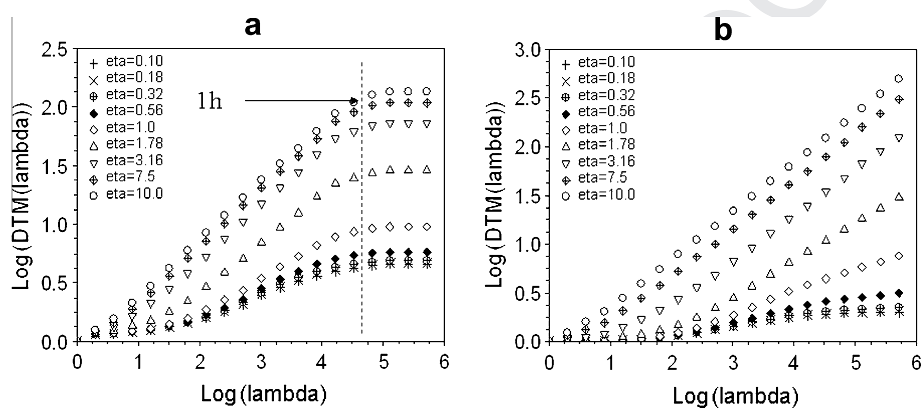

Fig. 12. Log–log plot of the Double Trace Moments of Saint Andre de Roquepertuis (a) and Orgeval (b). The latter displays a rather clear multifractal behaviour, whereas the former displays a scaling break at about 1 h. In both cases, these curves exhibit a multiscaling behaviour similar to their TM counterparts, i.e. respectively Figs. 8b and 7b.

The particular usefulness of the DTM method is that its scaling function $K(q,\eta)$ (with $K(q,1) = K(q)$) for the universal multifractals satisfies the following relationship:

$$K(q,\eta) = \eta^{\alpha}K(q) \tag{5}$$

which enables to estimate the index of multifractality α with the help of the slope of $K(q,\eta)$ vs. η in a log–log plot. C_1 is then estimated with the help of the linear fit in the log–log plot (e.g. its intersection with the axis $\mathrm{Log}(\eta) = 1$) and Eq. (3). However, as illustrated by Fig 11b, due to the finite size of the samples, as well as to detectability threshold, this relation (Eq. (5)) holds only over a finite range $[\eta_{\min}, \eta_{\max}]$ of η values. Following (Hittinger, 2008), these bounds can be theoretically estimated with the help of the corresponding second order phase transitions (Schertzer and Lovejoy, 1992) as being:

$$\eta_{\min} = (c_{\Sigma}/C_1)^{1/\alpha}\max(1, 1/q) \tag{6}$$

$$\eta_{\max} = (d/C_1)^{1/\alpha}\min(1, 1/q) \tag{7}$$

where c_{Σ} is the codimension of the empirical rainfall support Σ, i.e. when the rain rate is estimated as non-zero, which can be estimated with the help of a box-counting algorithm and depends on the detectability threshold, whereas d is the dimension of the embedding space ($d = 1$ for time series). Nevertheless, the main difficulty with Eqs. (6) and (7) is their nonlinear dependence on α, i.e. you need a first estimate of α to examine over which range $[\eta_{\min}, \eta_{\max}]$ you can obtain a good estimation of α. Therefore, one may first choose an intermediate value $\bar{\eta}$ to compute in its neighbourhood a first guess of the slope of $K(q,\eta)$ vs. $\mathrm{Log}(\eta)$, e.g. with the help of a given number n (usually $n = 6$) of the nearest discretised η values. It is rather convenient to define $\bar{\eta}$ by:

$$K(q,\overline{\eta}) = (K(q,\eta)_{\min}K(q,\eta)_{\max})^{1/2} \tag{8}$$

where $K(q,\eta)_{\min}$ and $K(q,\eta)_{\max}$ correspond respectively to the lower and upper empirical values of $K(q,\eta)$, i.e. the two horizontal branches of this curve (see Fig. 11b), and therefore to rough estimates of $K(q,\eta_{\min})$ and $K(q,\eta_{\max})$.

Within the first estimate of $[\eta_{\min}, \eta_{\max}]$, we developed two different, new variations of the DTM algorithm, in order to verify whether one of them is significantly less sensitive to the data quality.

The first variation is the DTM algorithm with an inflection point (DTM-IP): one looks for the inflection point $\mathrm{Log}(K(q,\eta_{IP}))$ of $\mathrm{Log}(K(q,\eta))$ vs. $\mathrm{Log}(\eta)$ inside of the range $[\eta_{\min}, \eta_{\max}]$, and estimate α with the help of the slope at this point (using again a given number m (e.g. $m = 6$) of the nearest discretised η values).

Poor quality data having no clear scaling behaviour will exhibit numerous inflection points and therefore generate a large dispersion of parameter estimates. In order to reduce these uncertainties, we developed the iterative procedure. This method is based on the idea that one may obtain by iteration a better precision on the η range over which α should be estimated. Hence, the second variation is the DTM method with a reduced range of η values (DTM-RR): the DTM-IP procedure is used for a second approximation of (α, C_1) in order to estimate $[\eta_{\min}, \eta_{\max}]$ with the help of Eqs. (7) and (8). Then these (α, C_1) are re-estimated with the help of a linear regression of $\mathrm{Log}(K(q,\eta))$ vs. $\mathrm{Log}(\eta)$ over $[\eta_{\min}, \eta_{\max}]$.

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

While the test results of these procedures on synthetic multifractal fields will be discussed elsewhere, in Section 5 we discuss an adapted Nash criterion to compare the respective DTM-IP and DTM-RR multifractal estimates obtained on the MF-5P database.

The obtained estimates confirm that, independently of the type of assessment procedure, the question of the data quality persists. Fig. 12 with no surprise illustrates that the DTM curves display spurious scaling breaks similar to those of the TM curves. The scaling breaks at about 1 h, e.g. the time series of Nîmes (Fig. 8a) and Saint Andre de Roqueqertuis (Figs. 8b and 12a), could be explained by the deficit of high frequency rainfall episodes. Indeed these series have only an hourly time resolution, hence a visible deficit of "homogeneous" episodes of shorter durations.

This means that the range of durations over which the universal multifractal parameters can be safely estimated is very sensitive to the quality of high frequency data. Not taking care of this question may lead to a serious increase of uncertainties in the hydrological estimates, e.g. in the Depth–Duration–Frequency (DDF) curves that are widely used in the hydrology.

Fig. 13 displays the DDF curves of the annual maxima for the Orgeval time series. The measured data having a 5 min resolution were also up-scaled to hourly data. Hence, the curves represented by round and square symbols correspond respectively to 5 min and hourly durations. Then the curve represented by triangle symbols for 5 min duration is obtained with the help of uniform disaggregation of hourly data, i.e. a transformation of hourly to 5 min precipitations were obtained by uniformly distributing the rainfall over the 12 time sub-intervals of 5 min. This illustrates the effect of using the hourly data to artificially construct the DDF curves for sub-hourly durations. For the duration of 5 min, the Fig. 13 shows a significant difference between the DDF curves corresponding to the "true" 5-min data and the corresponding artificial data.

Due to lack of small scale variability, the latter, obtained with the help of a uniform disaggregation from hourly rainfall data (triangles) remain weaker than those estimated from the observed precipitations (circles). For instance, the depth for a return period of 15 years (and a 5 min duration) drastically decreases from 22 to 4 mm. This illustrates that the deficit of "homogeneous" episodes of shorter duration would cause important underestimates of the rainfall for operational applications in the hydrology. We will pursue elsewhere the analysis of these hydrological consequences with the help of rainfall–runoff models.

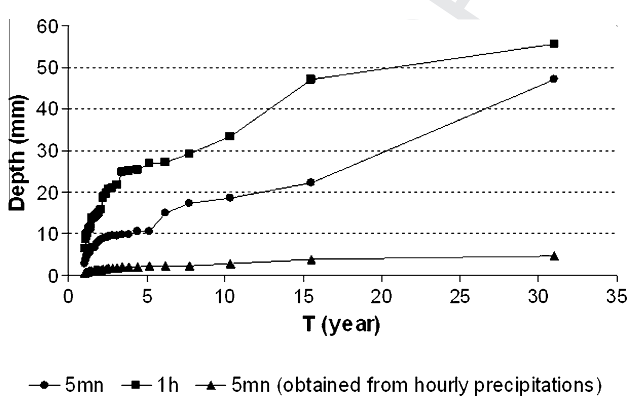

Fig. 13. The DDF curves of the annual maxima of the Orgeval time series. Square symbols correspond to hourly duration. For 5 min duration corresponding to round and triangle symbols, there is a significant difference between the DDF curve obtained from the observed precipitations (rounds) and the one obtained with the help of uniform disaggregation from the hourly precipitations (triangles). This confirms the need to well assessed the effective frequency of measurement and record.

4. SERQUAL: a procedure to select high quality data sequences

The observed sensitivity of the results of scaling analysis to small scale data quality lead us to develop an automatic procedure SERQUAL, written in the open access programming language SCILAB (Pinçon, 2000), that allows to quantify the quality of the time series not only over the whole time series, but also period by period, e.g. in this paper we started from year by year analysis, although the method is not at all limited to this period choice. Indeed, the yearly period was considered in the present paper due to the somewhat surprising observation that indeed the quality of the time series is in general far from being uniform and even monotonous, e.g. the data quality can decline in most recent years! However, for dryer climates a longer period might be needed.

The procedure SERQUAL is based on the conjunction of three criteria. The first one is the observation quality, which is measured by the percentage of missing data as follows:

$$r_{mis} = \frac{T_{mis}}{T_{tot}} \tag{9}$$

where T_{mis} is the total time length of missing data and T_{tot} is the total length of the data record.

The second criterion is the quality of the effective time resolution. It is measured by the episode duration having the highest probability, where the probability of the δ_i episode duration is estimated as follows:

$$\Pr(\delta_i) = \frac{N(\delta_i)}{\sum_j N(\delta_j)} \tag{10}$$

where $N(\delta_i)$ is the number of the episodes of duration δ_i. Therefore, when the probability of the 5 min episode duration is the highest one, the data sequences are considered having the effective 5 min resolution. At first glance, it is surprising to note that the effective resolution may change depending on the period of year for the same station. Let us remember that a detailed transcription from the paper record graphs was often performed only for some sever rainfalls. Therefore, these detailed small-scale data were inserted into rainfall records of a lower time resolution. These facts explain the observed changes in effective duration in the time series of a given station.

The final criterion is the quality of the probability distribution of episode durations. This quality is measured with the help of the determination coefficient (R^2) of the power law (over a given range of durations). In this paper, the determination coefficient (R^2) is estimated with the help of a linear regression of $\mathrm{Log}(\Pr(\delta))$ vs. $\mathrm{Log}(\delta)$ over the range of small durations from 10 min to 2 h 30 min. If the R^2 coefficient is higher than 0.8, the quality is graded as "A1"; if it ranges from 0.65 to 0.8, the quality is graded as "A2"; from 0.5 to 0.65, the quality is graded as "A3". Finally the quality is graded as "0" if the determination coefficient is less than 0.5.

One may note that the two last criteria require that SERQUAL should be only applied to long enough records. For each of these criteria, Fig. 14a–c displays geographical maps indicating the quality of the 166 rainfall time series from the MF-5P database. The corresponding data quality rates can be finally pooled together to get an overall data quality estimator.

Applying the SERQUAL procedure on the 166 rainfall time series, the results show (Fig. 14a) that most of the data have only a hourly resolution (110 series among 166, or 66%). There are only 40 rainfall time series having the effective resolution of 5 min. The corresponding rainfall gauges are mainly concentrated in three administrative regions of France: 15 gauges are in the region of Rhône-Alpes (Isère county with identification number 38), 6 gauges are in the region of Ile-de-France (Yvelines county with identification number 78) and 19 gauges are in the south of France

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

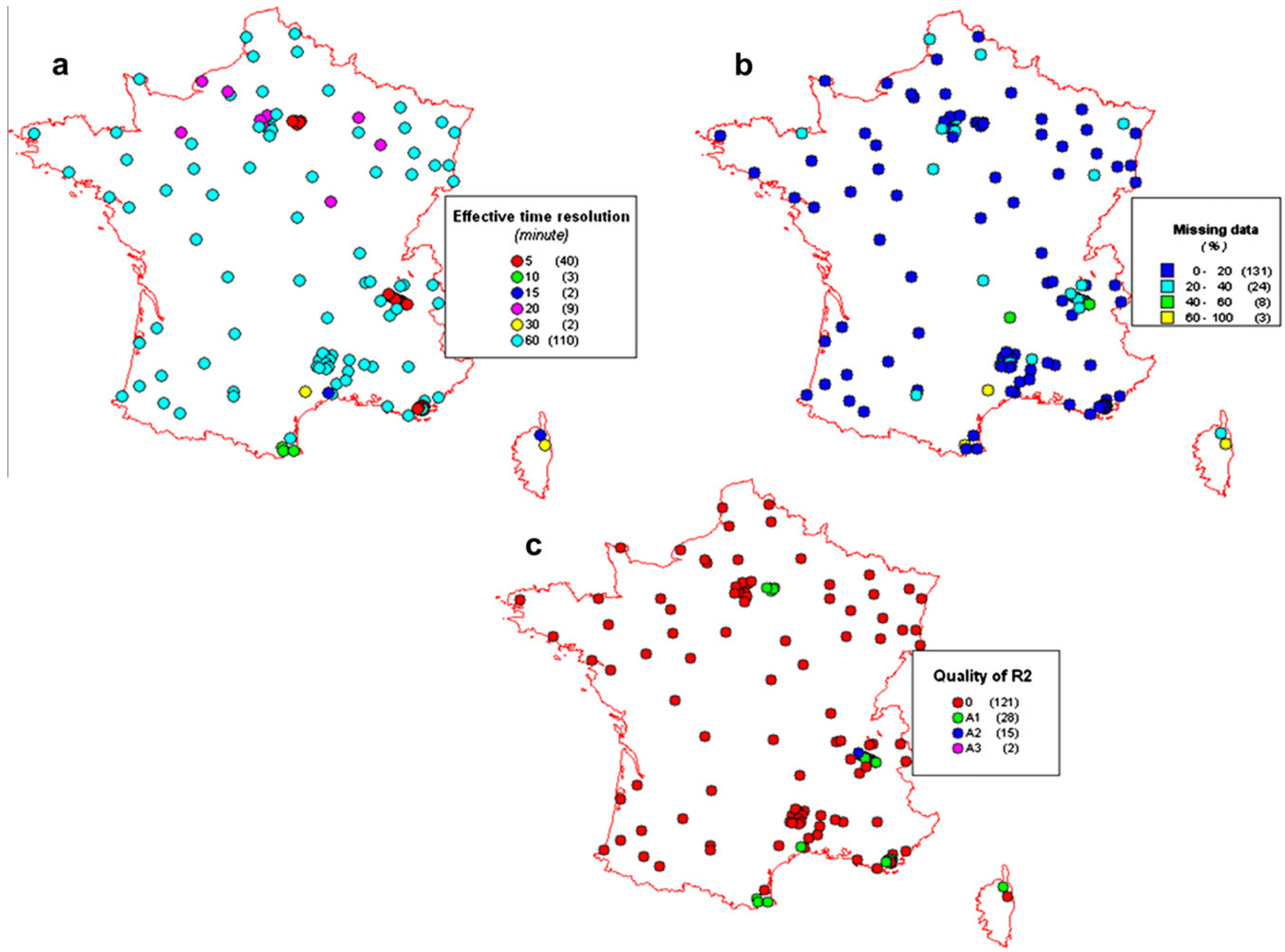

Fig. 14. Geographical maps of the quality of the series: effective time resolution quality (a: 5 min. (red), 10 min. (green), 15 min. (dark blue), 20 min. (purple), 30 min. (yellow), 60 min. (aqua)), missing data (b: 0–20% (dark blue), 20–40% (aqua), 40–60% (green), over 60% (yellow), probability distribution of episode durations (c: A1 for $R^2 \geqslant 0.80$ (green), A2 for $0.65 \leqslant R^2 < 0.80$ (dark blue), A3 for $0.50 \leqslant R^2 < 0.65$ (purple) and 0 for $R^2 < 0.50$ (red)). For these three figures, the corresponding number of stations corresponding to a given quality level with respect to a given criterion is displayed between parenthesis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Var county with identification number 83). It is important to note that when applied to year by year sequences, instead of doing it to the whole series, the SERQUAL procedure shows that all the series of counties 38 and 78 (respectively 15 and 6 series) have an effective time resolution of 5 min. On the contrary, the 19 series of the county 83 exhibit a much more complex behaviour due to a mixture of various effective time resolutions (Fig. 15). The data quality analysis also shows that among the 126 series that do not have an overall time resolution of 5 min, 39 series have a few (not necessarily consecutive!) years an effective time resolution of 5 min (Fig. 16), the other 87 time series do not have any year with such an effective time resolution.

Therefore, in order to perform our analysis on data having a uniform quality, we selected series or extracted sub-series that correspond to successive years having an effective time resolution of 5 min and for a length greater or equal to 5 years. According to this selection criterion, 15 time series were selected for the county 38 and 6 for the county 78. For the county 83, we could only extract subseries: 14 for the period 1989–2005, but only one for the periods 1987–2005, 1989–1995 and 1991–2005. Two series of this county did not have any of such periods.

Therefore, with the help of the SERQUAL procedure, only 38 time series, among 166 available ones, were selected as the high quality rainfall measurements, and in particular, as the full series having the effective data resolution of 5 min. These 38 time series have been used for further validation analysis discussed in the next section.

5. Discussion of the results

As we already mentioned, scaling methods in general remain very sensitive to data quality. For example, (Hittinger, 2008) discusses many difficulties in automatic routines of the DTM-IP method when looking for a precise estimate of the multifractal parameters α and C_1. To quantify the uncertainties of parameter estimates, we decided to compare the DTM-IP results with those obtained with the help of the iterative procedure DTM-RR. In both cases, we started with rather crude estimates of α and C_1, which analytically define the maximum probable singularity γ_s of a sample and therefore the upper bound of statistical orders q and η that can be used in the TM and DTM analyses. The determination coefficient R^2 can be furthermore used as an indicator of quality for linear fits. While, the iterative methods become much less sensitive to the crudity of initial estimates and the quality of original data, they do not distinguish the real data breaks from the spurious ones.

We decided to apply both DTM-IP and DTM-RR methods to the 38 time series of 5 min rainfall that were pre-selected by the SER-

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044

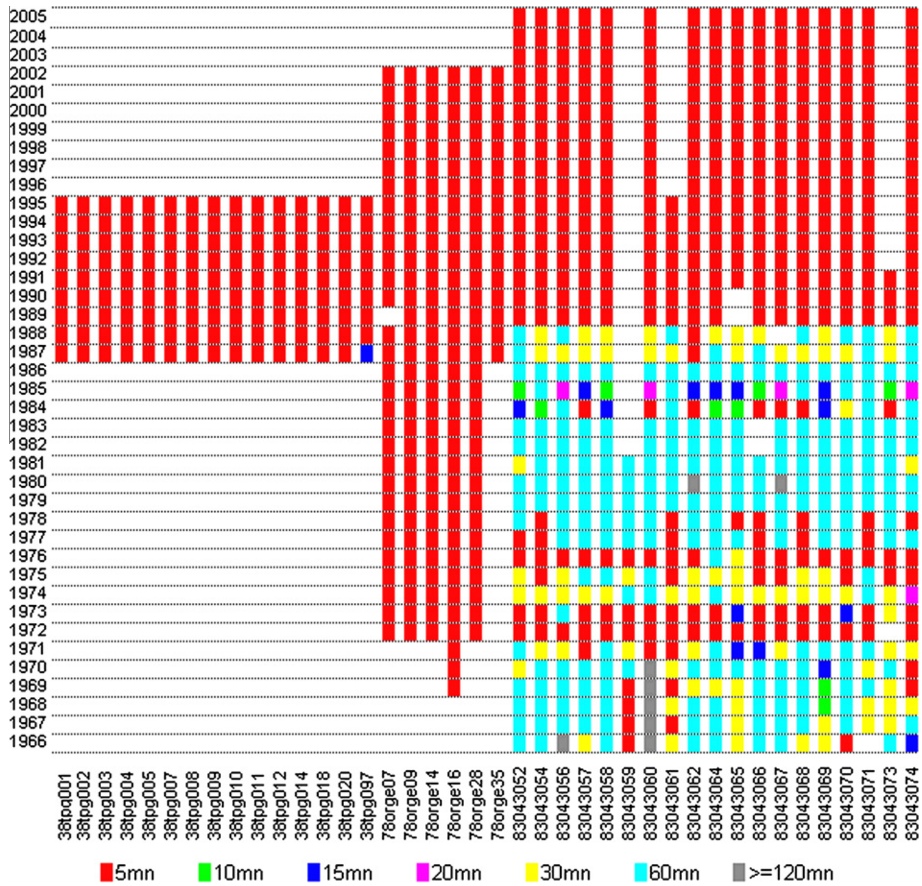

Fig. 15. Effective time resolution quality (ranging from 5 min to less than 120 min, see colour code) estimated year by year of the 40 series having an overall 5 min resolution. The displayed effective time resolution is rather inhomogeneous and non stationary. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

QUAL procedure. We intended to compare the multifractal parameter estimates, obtained by these two methods: close estimates would support the idea that the SERQUAL criteria assure a sufficient data quality in order to detect an unperturbed scaling behaviour and over the appropriate range of scales without the necessity of iterative methods. In order to compare the parameter estimates obtained by the two methods, we use a Nash criterion adapted to our situation.

The Nash coefficient is defined in the following manner

$$Nash = 1 - \frac{\sum (Y_i - X_i)^2}{\sum (X_i - \bar{X})^2} \tag{11}$$

where the X_i are the reference values (usually, the observations), $\bar{X}$ their average, and Y_i the predicted values. This coefficient is bounded above by 1, which corresponds to a perfect prediction, whereas lower and lower values correspond to worst and worst prediction. For instance, 0 corresponds to a prediction that is not better that the average $\bar{X}$.

Because the multifractal parameters obtained by the DTM-RR method are much less sensitive to the data quality, we will use them as the reference data with respect to the classical Nash criterion, whereas the multifractal parameters obtained by the DTM-IP would be used instead of the model prediction. Fig. 17 displays good correspondences between the estimated multifractal parameters that agree with the Nash parameters of 0.85 for α-estimates (Fig. 17a) and of 0.96 for C_1-estimates (Fig. 17b).

We have now to discuss another feature of the high resolution rainfall data we used. Meteo-France is in charge of maintaining most of the archives of the precipitation data in France and generally have been using a 6 min time step for short duration rainfall (the MF-6P database), whereas in the above sections we primarily discussed the results obtained on the 5 min rainfall (MF-5P database). It was therefore instructive to compare these two databases, in particular by performing quality tests for the same stations over the same periods.

Table 1 displays the main features of this comparison on the example of 3 stations: Brest, Mont Aigoual and Marseille. It turns out that the effective time resolution is higher for the MF-6P time series, whereas the rain period is longer for the MF-5P time series, in particular for the Marseille series over the period 1982–1988.

Among the rainfall data of Marseille, with the help of the SERQUAL procedure we finally selected 2 years having the best data quality, i.e., the year 1986 with an effective 5 min resolution (from MF-P5) and the year 1999 with an effective 6 min resolution (from MF-P6). Fig. 18 displays the corresponding probability distributions of the durations of « homogeneous » rainfall episodes during each year. This figure puts on evidence the existence of the power law with an exponent close to 2 that stands up to the highest resolution of 6 min rainfall. Indeed, while a 5 min resolution remains dominant for the year 1986, a power law could be roughly fitted up to 10 min resolution, which was the one selected for the SERQUAL procedure. Therefore, strengthening the third criterion of the SERQUAL procedure that evaluates the quality of probability distribution, i.e. estimating the power law over the full range of small durations, would lead to the rejection of the 5 min Marseille data. In this paper we gave preference to a looser criterion (i.e., with 10 min resolution) in conjunction with the criterion on the effec-

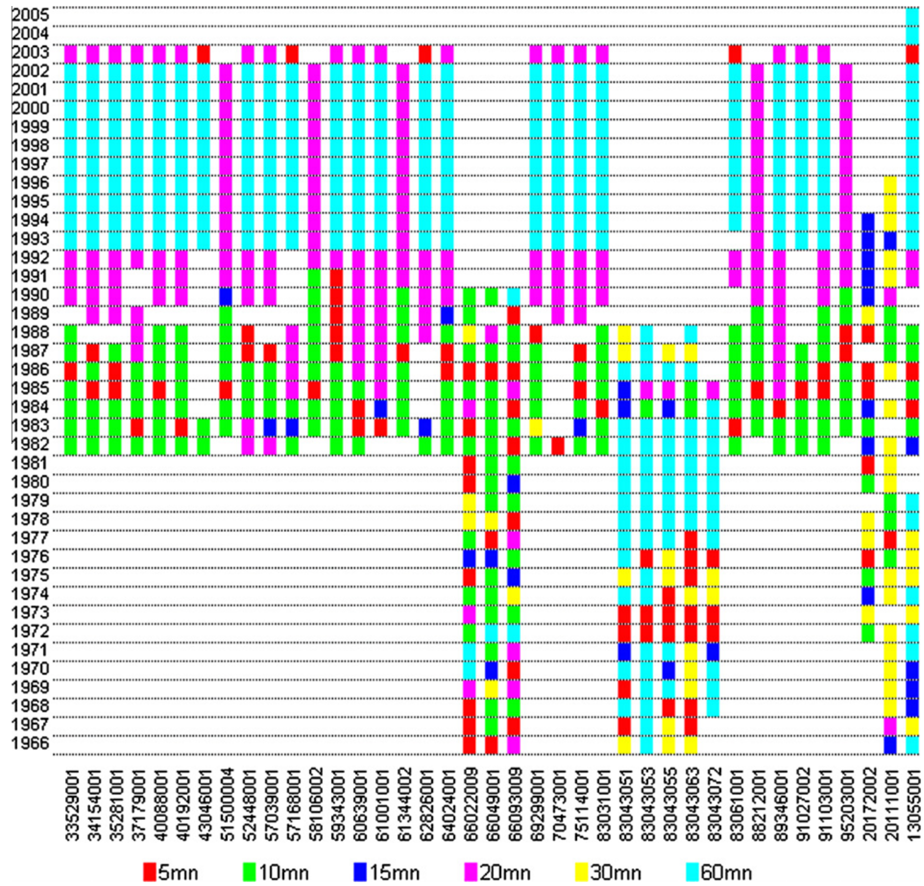

Fig. 16. Effective time resolution quality year by year of the 39 series which do not have an overall effective time resolution of 5 min, but have at least 1 year of 5 min resolution. The displayed effective time resolution is even more inhomogeneous and non stationary than that of Fig. 15.

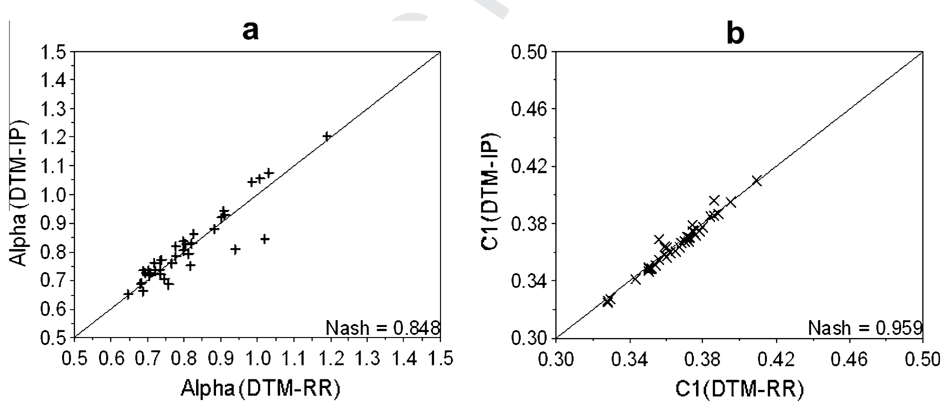

Fig. 17. Correlation of parameters α and C_1 estimated respectively by the DTM-IP and DTM-RR methods, as well as the corresponding Nash coefficient.

Table 1
Intercomparison between the MF-5P and MF-6P databases.

Station	Brest		Mont_Aigoual		Marseille			
Period	1990–2003		1992–2003		1997–2004		1982–1988	
Database	5P	6P	5P	6P	5P	6P	5P	6P
Effective time resolution	Hourly	18 min	Hourly	18 min	Hourly	6 min	10 min	6 min
Missing data duration/period duration (%)	11.3	9.2	9.0	5.3	3.5	1.8	3.3	3.3
Rainfall duration/period duration (%)	13.4	6.8	12.3	7.8	4.1	1.71	19	3.3

tive time resolution quality. Let us anyway mention that such a departure from the power law, which has been observed for the highest resolution, could be related to a rather unexpected increase of the fractal dimension of the rainfall support. Indeed, Fig. 19 displays the fractal dimensions of the rainfall support in Marseille during the years 1986 and 1999. With no surprise, the rainfall support fractal dimension at lower resolution converges to the dimension of the embedding space: it always rains at monthly scales. For

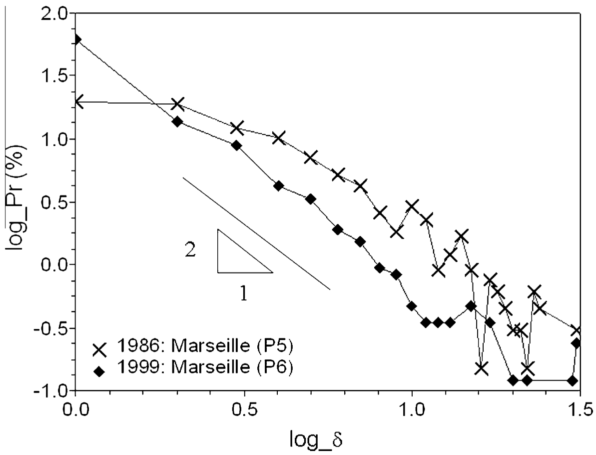

Fig. 18. Comparison (in a log–log plot) of the duration probability distribution of the Marseille homogeneous episodes during the year 1986 (crosses) and 1999 (filled squares) corresponding respectively to the year having the best data quality for MF-P5 and MF-P6. The existence of a power law up to the highest resolution is evident only for the MF-P6 data. The reference straight line has a slope 2.

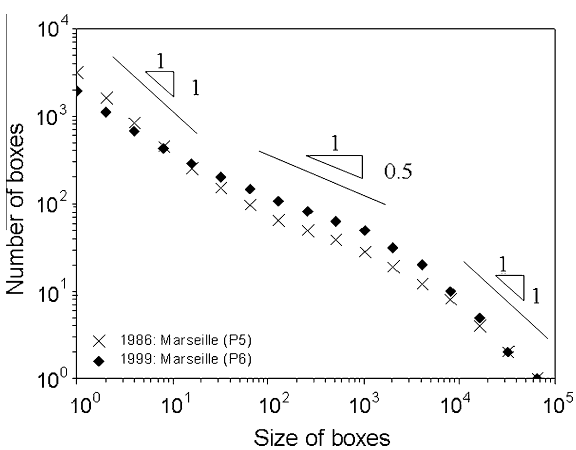

Fig. 19. Comparison (in a log–log plot) of the rainfall support dimension in Marseille during the year 1986 (crosses) and 1999 (filled squares) corresponding respectively to the year having the best data quality for MF-P5 and MF-P6. For MF-P5, the fractal dimension of sub-hourly rainfall converges to the dimension of the embedding space. For convenience, the reference straight lines of slopes 1, 0.5 and 1 are displayed.

higher time resolutions (from a month to about an hour) the rainfall is known to be an intermittent process and indeed the support fractal dimension is of the order of 0.5. However, instead of having the same intermittent behaviour for sub-hourly durations, there is an increase of the support fractal dimension for both rainfall data of Marseille. This is more obvious for the 5 min resolution data, where the fractal dimension reaches again the embedding space dimension. Such a behaviour suggests that a given downscaling algorithm was applied to hourly data to obtain a seemingly 5 min data resolution and which introduces spurious uniform frequencies of data recording. Since such a behaviour was observed for other stations, future research on high resolution data sets is needed to clarify this issue.

6. Conclusions

The obtained results show that the deficit of high frequency episodes causes scaling breaks and therefore difficulties to define scaling laws of the rainfall with many consequences for operational hydrology. Indeed, whereas scaling methods provide powerful tools for nowadays hydrology, we demonstrated that a particular attention should be paid to the data quality in order to adequately interpret the often observed scaling breaks and their consequences on the parameter estimates.

In this direction, we developed a first version of a SERQUAL procedure to automatically detect the effective time resolution of highly mixed data. Being applied to the 166 rainfall time series (MF-P5 database), the SERQUAL procedure has detected that most of them have an effective hourly resolution, rather than a 5 min resolution. Furthermore, series having an overall 5 min resolution do not have it for all years. Indeed the effective resolution may have been unstationnary and changed for a given station. This is because the majority of long time series correspond to continuous records of hourly rainfall, into which sub-hourly data had been inserted based on the transcription from the paper record graphs. Unfortunately such transcriptions were often performed only for some periods of time merely corresponding to sever rainfall events. The systematisation of automatic recording has eliminated this problem for the recent years records, but not for the older parts of the time series. At first, this raises serious concerns on how to benchmark stochastic rainfall models at a sub-hourly resolution, which are particularly desirable for operational hydrology. Therefore, database quality must be checked before use. Furthermore, we showed that our procedure SERQUAL enable us to extract high quality sub-series from longer time series that will be much more reliable to calibrate and/or validate short duration quantiles and hydrological models.

Acknowledgements

We thank A. Bardossy (Editor), E. Kahya (Associate Editor) and two anonymous Referees for their valuable comments and suggestions that greatly helped us to improve our manuscript. We highly acknowledge enlightening discussions with Ph. Dandin and J.M. Veysseire who also facilitated the access to the rainfall data of Meteo-France in the framework of the project "Multiplicity of Scales in Hydrology and Meteorology (Meteo-France – Ecole des Ponts ParisTech). Partial financial supports from the Regional Research Network on Sustainable Development R2DS (project GARP-3C) and from the Chair "Hydrology for Resilient Cities" (sponsored by Veolia) of Ecole des Ponts are highly acknowledged.

References

Arnaud, P., Lavabre, J., 1999. Using a stochastic model for generating hourly hyetographs to study extreme rainfalls. Hydrol. Sci. J. 44 (3), 433–446.
Aronica, G.T., Freni, G., 2005. Estimation of sub-hourly DDF curves using scaling properties of hourly and sub-hourly data at partially gauged site. Atmos. Res. 77 (1–4), 114–123.
Berggren, K., 2007. Urban Drainage and Climate Change—Impact Assessment. Luleå University of Technology.
Berndtsson, R., Niemczynowicz, J., 1988. Spatial and temporal scales in rainfall analysis – some aspects and future perspectives. J. Hydrol. 100 (1–3), 293–313.
Berne, A., Delrieu, G., Creutin, J.D., Obled, C., 2004. Temporal and spatial resolution of rainfall measurements required for urban hydrology. J. Hydrol. 299 (3–4), 166–179.
de Lima, M.I.P., 1998. Multifractals and the Temporal Structure of Rainfall. Ph.D. Thesis, Wageningen Agricultural University, Wageningen, 229 pp.
Fankhauser, R., 1997. Measurement properties of tipping bucket rain gauges and their influence on urban runoff simulation. Water Sci. Technol. 36 (8–9), 7–12.
Fankhauser, R., 1998. Influence of systematic errors from tipping bucket rain guages on recorded rainfall data. Water Sci. Technol. 37 (11), 121–129.
Hittinger, F., 2008. Intercomparaison des incertitudes dans l'Analyse de Fréquence de Crues classique et l'Analyse Mutlifractale de Fréquence de Crues. Ing. Dipl. Thesis, Institut National Polytechnique de Grenoble, Grenoble, 35 pp pp.
Hoang, C.T., 2008. Analyse fréquentielle classique et multifractale des 10 series pluviométriques à haute résolution. M.Sc. Thesis, Université P. & M. Curie, Paris, 50 pp.
Kolmogorov, A.N., 1941. Local structure of turbulence in an incompressible liquid for very large Raynolds numbers. Proc. Acad. Sci. URSS, Geochem. Sect. 30, 299–303.

Kundu, P.K., Bell, T.L., 2006. Space-time scaling behavior of rain statistics in a stochastic fractional diffusion model. J. Hydrol. 322 (1–4), 49–58.

Kvicera, V., Fiser, O., Riva, C. and Sharma, P., 2005. Comparison of Tipping-bucket raingauge record processing at various workplaces. In: Contribution PM9103 of the 3rd (Final) Workshop of the COST280 Propagation Impairment Mitigation for Millimetre Wave Radio, Phara, Czech Republic, pp 1–4.

LaBarbera, P., Lanza, L.G., Stagi, L., 2002. Tipping bucket mechanical errors and their influence on rainfall statistics and extremes. Water Sci. Technol. 45 (2), 1–9.

Lavallée, D., Lovejoy, S., Schertzer, D., Schmitt, F., 1992. On the determination of universal multifractal parameters in turbulence. In: Moffat, K., Tabor, M., Zalslavsky, G. (Eds.), Topological Aspects of the Dynamics of Fluids and Plasmas. Kluwer, pp. 463–478.

Niemczynowicz, J., 1999. Urban hydrology and water management – present and future challenges. Urban Water 1 (1), 1–14.

Obukhov, A.M., 1941. On the distribution of energy in the spectrum of turbulent flow, Izvestiya. Geogr. Geophys. 5, 453–466.

Ogden, F.L. et al., 2000. Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. J. Hydrol. 228 (1–2), 82–100.

Olofsson, M., 2007. Climate Change and Urban Drainage—Future Precipitation and Hydraulic Impact. Luleå University of Technology, 20 pp.

Pinçon, B., 2000. Une introduction à Scilab. Ecole National des Ponts ParisTech, Marne-la-Vallée.

Schertzer, D., Lovejoy, S., 1987a. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. J. Geophys. Res. D 8 (8), 9693–9714.

Schertzer, D., Lovejoy, S., 1987b. Singularités anysotropes, divergence des moments en turbulence. Ann. Sosiété mathématique du Québec 11 (1), 139–181.

Schertzer, D., Lovejoy, S., 1992. Hard and soft multifractal processes. Physica A 185 (1–4), 187–194.

Schertzer, D., Lovejoy, S., 1997. Universal multifractals do exist! J. Appl. Meteorol. 36, 1296–1303.

Schilling, W., 1991. Rainfall data for urban hydrology: what do we need? Atmos. Res. 27 (1–3), 5–21.

Schmitt, F., Lavallée, D., Schertzer, D., Lovejoy, S., 1992. Empirical determination of universal multifractal exponents in turbulent velocity fields. Phys. Rev. Lett. 68, 305–308.

Tchiguirinskaia, I., Bonnel, M., Hubert, P. (Eds.), 2004. Scales in Hydrology and Water Management, vol. 287. IAHS, Wallingford, UK, 170 pp.

Tessier, Y., Lovejoy, S., Schertzer, D., 1993. Universal multifractals: theory and observations for rain and clouds. J. Appl. Meteorol. 32 (2), 223–250.

Please cite this article in press as: Hoang, C.T., et al. Assessing the high frequency quality of long rainfall series. J. Hydrol. (2012), doi:10.1016/j.jhydrol.2012.01.044