Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off meeting held in Brasilia, 20–22 November 2002). IAHS Publ. 309, 2007.

Potential of multifractal modelling of ungauged basins

IOULIA TCHIGUIRINSKAI[1,2], DANIEL SCHERTZER[3], PIERRE HUBERT[2], HOCINE BENDJOUDI[2] & SHAUN LOVEJOY[4]

1 *CEREVE, Ecole Nationale des Ponts et Chausées, Marne-la-Vallee, France*
ioulia@cereve.enpc.fr
2 *UMR Sisyphe, LGA, Université Paris VI, Paris, France*
3 *CEREVE-ENPC, 6–8 ave Blaise Pascal 77455, Marne-la-Vallée Cedex, France*
4 *Physics Dept., McGill University, Montreal, Quebec, Canada*

Abstract River runoff phenomena reflect many complex interactions between diverse basin factors, as well as meteorological and climatic fluctuations that strongly modify the precipitation input. Therefore river runoff prediction and forecasting remains a rather tricky issue in hydrology. The question is even more puzzling for the areas where no actual flow measurements are available. Present communication concretely demonstrates how the multifractals help to overcome some present difficulties in modelling, predicting and forecasting river runoffs.

Key words multifractals; pan-Artic drainage region; R-ArticNET

INTRODUCTION

To ensure valuable results from multifractal analysis, we used about three million river runoff measurements at very different locations all over the world. The time-scale of these monthly data varies for different stations from several years to several centuries.

For basins ranging from 1 km^2 to more than 10×10^6 km^2, we found that spatial variability of the discharge is dominated by the spatial variability of the corresponding drainage areas. This result would appear trivial, but it is much more than a single correlation. Not only are river discharges and drainage areas multifractal over the full range of scales, but also they have the same index of multifractality. We present some consequences of these results, which are well beyond the scope of earlier fractal drainage area studies, e.g. the power-law of the probability distribution tail for the specific discharges.

Combining these new results with our earlier results on multifractal time variability, we address the question of space–time multifractal runoff analysis and modelling. The latter opens a way to multifractal predictions in ungauged basins.

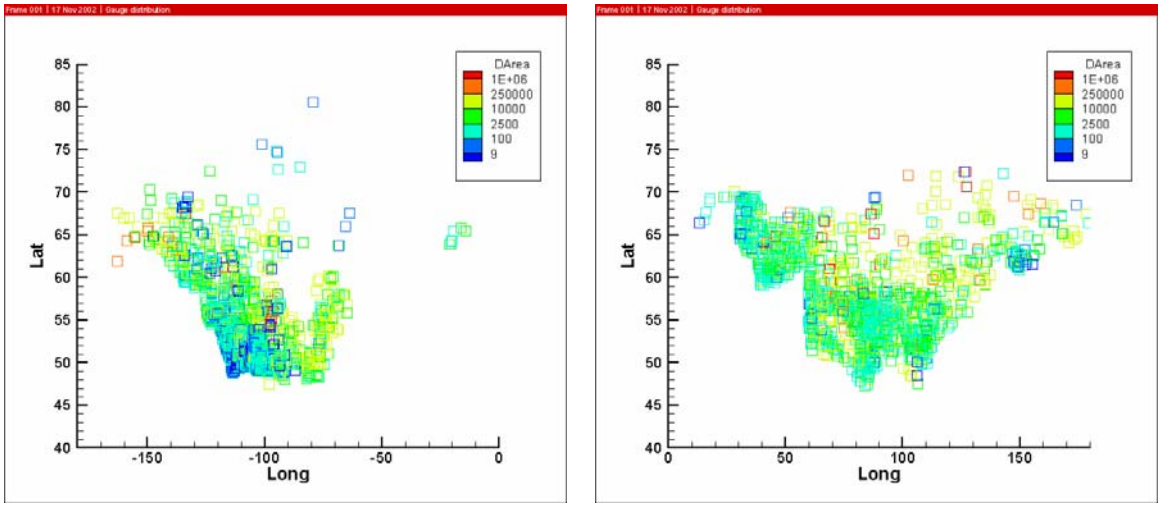
DATA DESCRIPTION

To establish the results of multifractal analysis and modelling, we used the second version of a web-based, hydrographic data network for the Arctic Region (R-ArcticNET v2.0). This is the comprehensive monthly river discharge database for the entire pan-

Arctic drainage system. The pan-Arctic drainage region represents a land area of more than 20×10^6 km^2 that is defined as all the terrestrial land area draining into the Arctic Ocean as well as the drainage regions of the Hudson Bay, James Bay and the Northern Bering Strait, which includes the Yukon and Anadyr Rivers. For the total of 3713 gauges that compose the R-ArcticNET database, the size of drainage areas ranges from 1 km^2 to more than 10^6 km^2. The number of gauges varies for each of the 11 hydrological zones of the pan-Arctic drainage region: Anadyr and Kolyma, 81; Lena, 199; Mackenzie, 500; Nelson, 1319; North European Russia, 319; Northwest Hudson Bay, 99; Ob, 603; Scandinavia, 15; South and East Hudson Bay, 183; Yenisei, 299; Yukon, 96 gauges. The source of the Canadian data was HYDAT (Environment Canada). Geographical location (as longitude/latitude coordinates) of the entire R-ArcticNET database and the corresponding distribution of drainage areas are presented in Fig. 1. This database contains monthly discharges from prior to 1900 (for four Canadian and five Russian gauges) until the early 1990s. The length of record for individual gauges is extremely variable and the majority of the data occurs between 1960 and 1990. Figure 2 displays not only a typical example of the intermittent character of the discharge time records, but also emphasizes the fact that the pan-Arctic drainage region rivers are completely icebound over the winter–spring period and that for most of them, the largest discharge input is due to snowmelt phenomena.

SHORT RECAPITULATION ON MULTIFRACTALS

Multifractals constitute a very convenient framework within which to analyse and simulate intermittent fields. In general, a multifractal field is obtained by a cascade process, whose paradigm can be traced back to the famous poem of Richardson (1922). An elementary process is repeated scale by scale, randomly transmitting a fraction of a given flux (e.g. energy flux for fluid turbulence) from a parent structure to its children



(a) (b)

Fig. 1 Network of discharge gauges of the entire R-ArcticNET database in longitude/latitude coordinates. Colours of circles indicate the size of drainage areas in km^2.

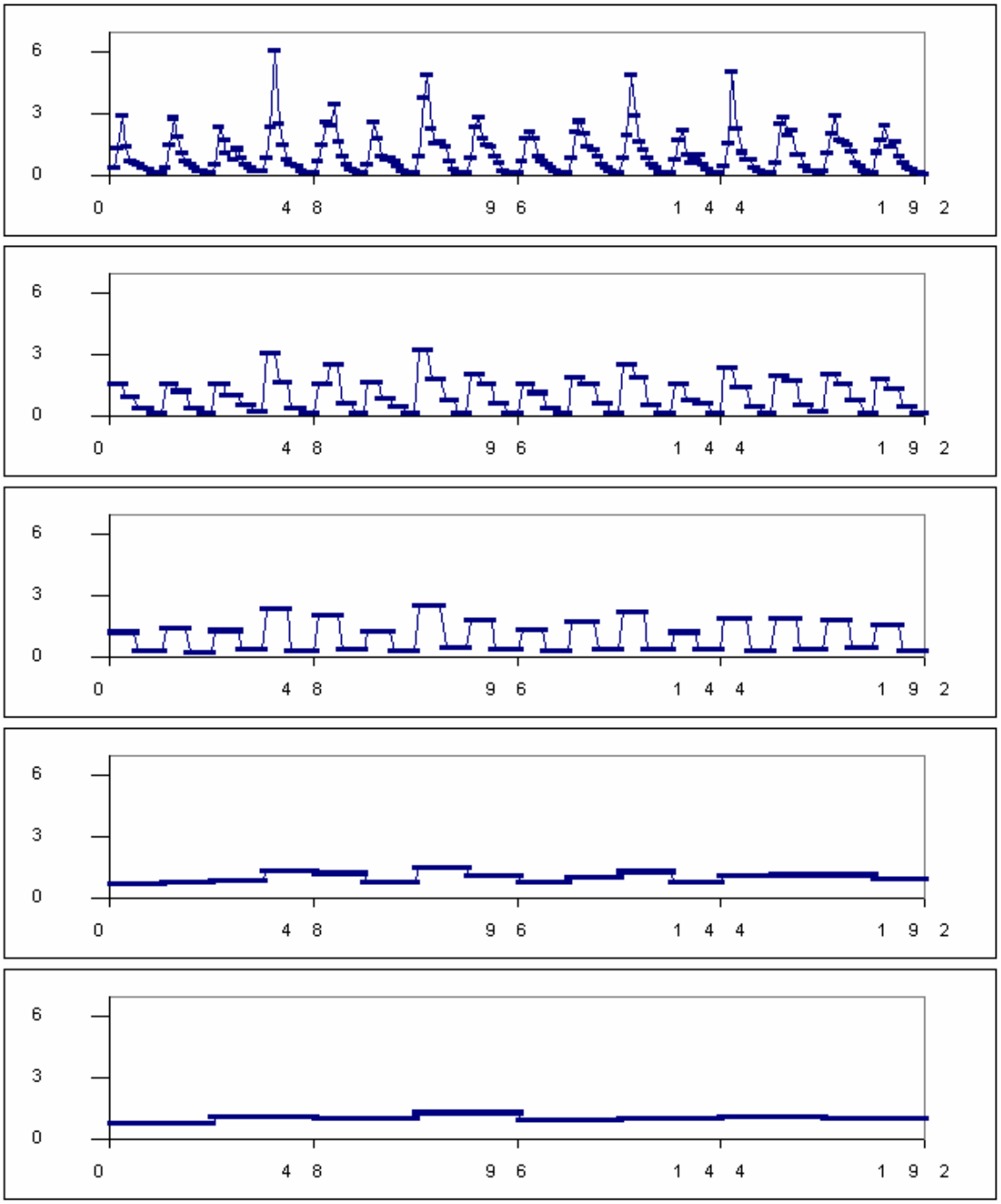


Fig. 2 16-year time series of elementary discharges at different time resolutions: 1 month, 3 months, 6 months, 1 year and 2 years (from top to bottom).

structures. In the simplest case, the scale ratio $\lambda = L/\ell$ (L is the external scale of the cascade, ℓ the scale of observation) takes discrete values: $\lambda = \lambda_1^n$ (λ_1 is the scale ratio for the discrete elementary step of the cascade, usually $\lambda_1 = 2$). The resulting flux, $F_{\lambda=2^n}(\underline{x}) = \prod_{i=1}^{n} \mu_i$, rather corresponds to the multiplication of identically independently

distributed (i.i.d.) random variables, μ_i (e.g. Monin & Yaglom, 1975; Schertzer *et al*., 1997). As $\lambda \to \infty$, the flux F_λ is no longer a point-wise function but a multi-singular measure. It means that the non-trivial limit $F \equiv F_\infty$ of the flux has values defined by any arbitrary small neighbourhood of a point $(\underline{x})$, but not by this precise point. So, this flux depends on the scale of the point neighbourhood. Usually, the flux over a given set should be conserved at all scales (e.g. its ensemble average should be strictly scale invariant in the case of a "canonical conservation").

The extreme spikes or the singularity of such flux may be defined in terms of a probability distribution by the order of singularity, $\gamma > 0$:

$$\Pr(F_\lambda \geq \lambda^\gamma) \approx \lambda^{-c(\gamma)} \tag{1}$$

where $c(\gamma)$ is the co-dimension function. The nonlinear behaviour of this function corresponds to the fact that high intensity events are less frequent than low intensity events, and therefore the most intense regions occupy a smaller fraction of the probability space, which means they have a larger co-dimension.

Therefore, a multifractal field corresponds to an infinite hierarchy of fractal sets and a definition of the co-dimension function in general requires an infinite number of parameters. Hence, let us focus particularly on the possibility of having so-called universal multifractals (Schertzer & Lovejoy, 1987). The very general idea of universality is that interactions between rather similar processes may lead them to converge to some attracting process. This fact has enormous consequences: only very few relevant parameters may define a stochastic process, whereas it could result from very complex interactions *a priori* requiring numerous parameters. It has been shown that by adding more and more intermediate levels in cascades or by multiplying i.i.d. cascades, one may reach a universal multifractal process.

PRINCIPALS OF THE UNIVERSAL MULTIFRACTAL ANALYSIS

For the universal multifractals, only three parameters H, α and C_1 are of fundamental importance. The first parameter indicates how the experimental data differ from a conservative flux. The latter could be directly modelled with the help of a multiplicative cascade. Fourier analysis is rather convenient to get a rough estimate of this parameter: the spectral slope $\beta = 2H + 1$ gives an estimate of $H < 1/2$ (Fig. 3). We may then apply the double trace moments technique (DTM; Lavallée *et al.*, 1992) directly to the river discharge data to obtain estimates of α and C_1. Otherwise, for accurate parameter estimates, the original data must be first passed through a filter that weights their Fourier components by k^H, with k being a wave number.

The main idea of DTM is to compute the ηth power of the original data (at their highest resolution) and to degrade this field at smaller and smaller resolutionλ. Then one may study the scaling behaviour of the statistical moments of order q of the resulting field $\varepsilon_\lambda^{(\eta)}$:

$$\left\langle \left(\varepsilon_\lambda^{(\eta)}\right)^q \right\rangle \propto \lambda^{K(q,\eta)} \tag{2}$$

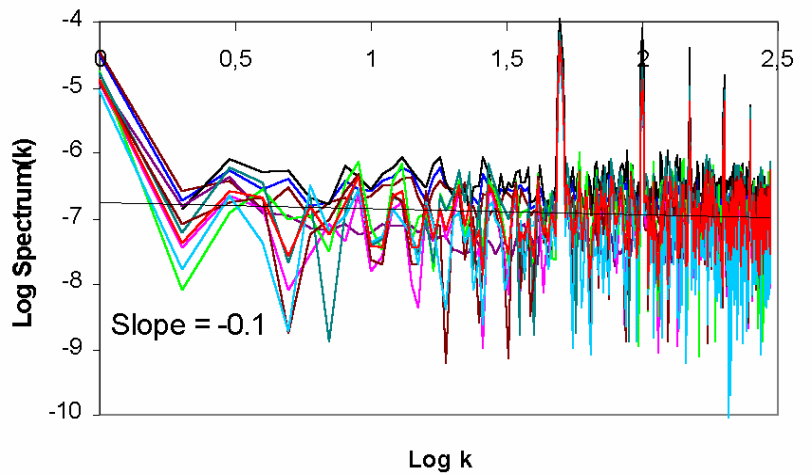


Fig. 3 Fifty-year spectra of monthly discharges representing all hydrological zones of the R-ArcticNET database which gave the first indication of the possible scaling behaviour of time series.

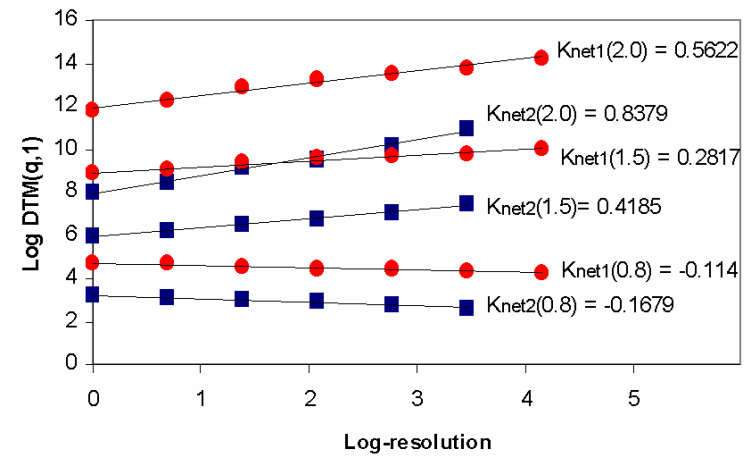


Fig. 4 Examples of a perfect scaling behaviour of the DTM curves for two-dimensional gauge networks analysis: Northwest Hudson Bay (blue squares) with 32 × 32 grid for subset of points from Fig. 1(a) (99 gauges) and Ob (red circles) with 64 × 64 grid for subset of points from Fig. 1(b) (603 gauges). Each slope gives an estimate of corresponding $K_{net}(q)$ for different q.

For universal multifractals, the scaling function of the double trace moments, $K(q,\eta)$, has a simple relation with the single moment scaling function ($K(q) \equiv K(q,1)$):

$$K(q,\eta) = \eta^{\alpha} K(q) \tag{3}$$

Equations (2) and (3) yield a convenient method for empirically defining the Lévy parameter α (α = 2 corresponds to the Gaussian case). Then the parameter C_1, which corresponds to the co-dimension of the mean singularity, may be computed according to the following equation for the universal multifractal scaling function (Schertzer & Lovejoy, 1987):

$$K(q) + Hq = C_1(q^{\alpha} \quad q)/(\alpha \quad 1) \tag{4}$$

Spatial data analysis: multifractality of discharges and corresponding drainage areas

The main difficulty of spatial data analysis can be seen in Fig. 1: the gauges have a notably inhomogeneous distribution in space. Therefore, it is impossible to attribute to these gauges a regular full-field two-dimensional grid of a sufficiently high resolution $\lambda \times \lambda$. The "sufficiently high" would be a resolution that allows study of the scaling behaviour of two-dimensional fields according to the DTM procedure described above. Therefore, in order to reach such resolution, we may accept that some of the grid cells will not include any gauges. But in order not to influence the DTM results by the artificial, non-measured zero values of a given measured field, one may consider that this field, φ_λ, results from the intersection of a full two-dimensional multifractal field, ε_λ, with a network of gauges, ρ_λ. This network of gauges is constructed on the principal that any grid cell with at least one gauge inside will have the value of 1, whereas the cells with no gauges inside will have the value of 0. The main consequence of such construction is that any ηth power of the network data would be identical to their 1st power: $\rho_\lambda^{(\eta)} = \rho_\lambda^{(1)}$. Since two fields ($\varepsilon_\lambda$ and ρ_λ) are statistically independent, using the DTM procedure with the help of equation (2) one may obtain:

$$<(\varphi_\lambda^{(\eta)})^q> = <(\varepsilon_\lambda^{(\eta)})^q><(\rho_\lambda^{(1)})^q> \quad \Rightarrow \quad K_{mes}(q,\eta) = K(q,\eta) + K_{net}(q) \tag{5}$$

Therefore, the true universal multifractal scaling function of the double trace moments, $K(q,\eta)$, would correspond to a difference between the scaling function of a given measured field and the scaling function of the network of gauges: $K(q,\eta) = \eta^\alpha K(q) = K_{mes}(q,\eta) \quad K_{net}(q)$. Figure 5 demonstrates the importance of such network correction for proper estimation of multifractal parameters and establishes a direct proof of the multifractal behaviour of the spatial distributions of

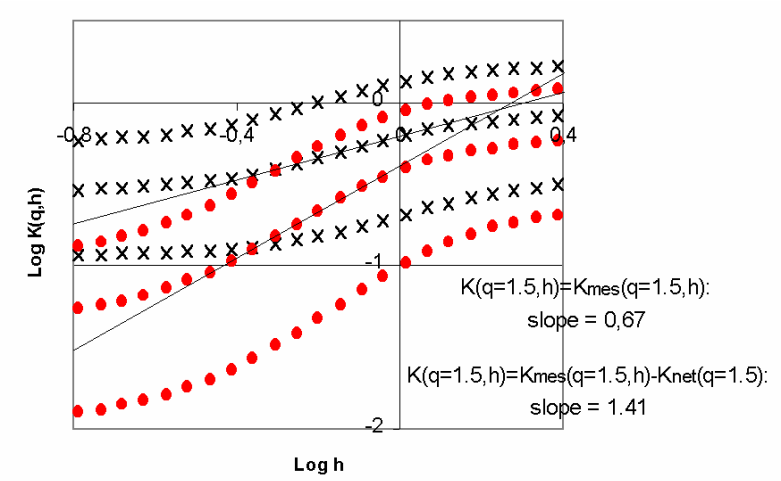


Fig. 5 Example of a log–log plot of empirical scaling function $K(q,\eta)$ *versus* η–moments for (from top to bottom) q = 2.0, 1.5 and 0.8 for drainage area measurements without spatial network correction (black crosses) and with network correction (red dotes). The straight lines indicate two slopes that correspond to α = 0.67 before correction and α = 1.41 after the network correction. For each q, the intersection of such curves with the axis logη = 0 gives the corresponding log$k(q)$. The universal multifractal expression for $K(q)$ was then used to compute C1 = 0.56 with corrected α.

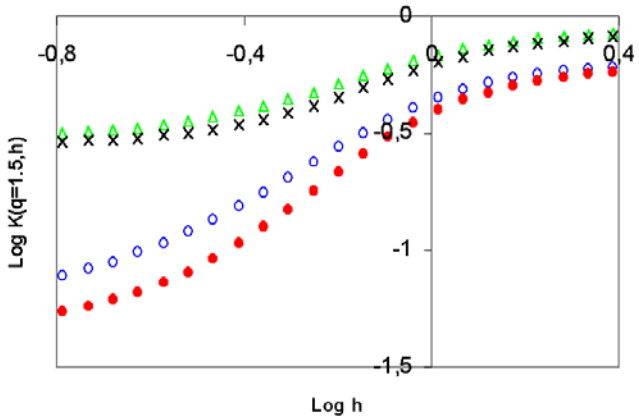


Fig. 6 A log–log plot of empirical scaling functions $K(q,\eta)$ *versus* η–moments for q = 1.5 from Fig. 5 in comparison with similar curves for spatial monthly discharge distribution (sampled during the year 1985): without spatial network correction (green triangles) and with network correction (blue circles). It corresponds to network corrected parameters $\alpha = 1.28$ and C1 = 0.8.

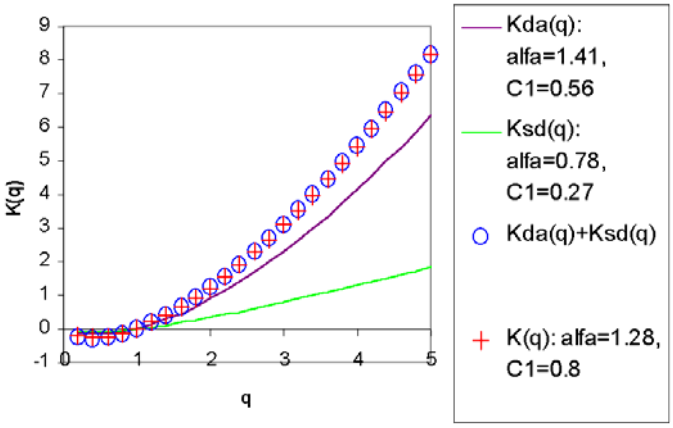


Fig. 7 Empirical (spatial analysis) scaling moment function $K(q)$ of monthly discharges (red crosses) compared with a sum (blue open dots) of empirical (space analysis) scaling moment function $K_{da}(q)$ of drainage areas (violet line) and of specific discharges, $K_{sd}(q)$ (green line).

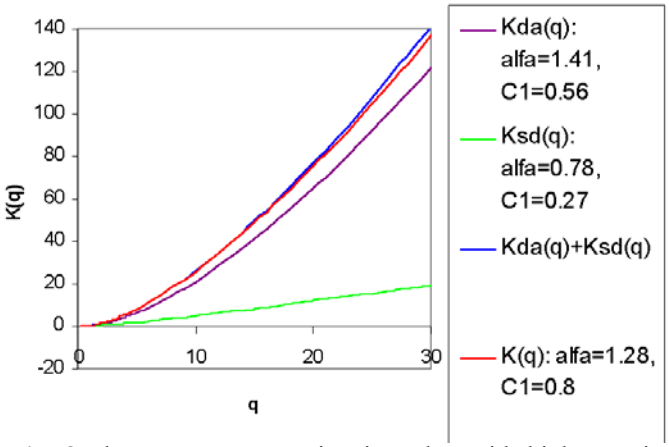


Fig. 8 The same curves as in Fig. 7 but with higher statistical moments. Empirical (spatial analysis) scaling moment function $K(q)$ of monthly discharges (red line) becomes distinct from the sum of two other scaling functions only at a statistical moment order of about 25.

monthly river discharges and of their drainage areas. To the authors' knowledge, this is the first time that the multifractality of drainage areas has been considered; it is also the first time that the multifractal analysis of river discharges was fulfilled in the spatial domain. Not only are discharges and drainage areas multifractal over the full range of scales but they also have rather similar indexes of multifractality (Fig. 6). These results are well beyond the scope of earlier fractal drainage area studies. Furthermore, multifractal space analysis has demonstrated (see Figs 7 and 8) that the sum of the empirical scaling moment function of drainage areas, $K\mathrm{da}(q)$, and of specific discharges, $K\mathrm{sd}(q)$, may be well approximated by an empirical estimate of the scaling moment function of monthly discharges up to a rather high order of statistical moments.

Time series analysis and modelling: S-shaped DTM curves and seasonal periodicity

Monthly discharge data systematically yield non-straight lines for DTM curves, i.e. $\log(\varepsilon_\lambda^{(\eta)})$ *versus* $\log(\lambda)$, instead of the expected straight lines corresponding to equation (2); an example is displayed in Fig. 9. This important fact was overlooked by Labat *et al.* (2002), Pandey *et al.* (1998) and Tessier *et al.* (1996), whereas it could be understood as a break of scaling. We will demonstrate that it corresponds to the seasonal periodicity, whose importance for river runoff phenomena in Russia has already been emphasized by Tchiguirinskaia *et al.* (2002), and calls for some generalization of the cascade processes. Indeed, the latter generally assume time translation invariance, which is incompatible with periodicity. Figure 10 displays perfect multiscaling behaviour for the range of scales selected from Fig. 9. On this range of scales we estimated the multifractal parameters as $\alpha = 1.56 \pm 0.3$ and $C1 = 0.28 \pm 0.08$, which may be used as input parameters for the time series simulation by a cascade process. The question we discuss below is how to introduce the seasonal periodicity for the synthetic discharges.

Another look at Fig. 2, which can be considered as a fragment of an inverted cascade process, helps us to understand how to introduce seasonal periodicity into a multiplicative cascade. First we can develop the usual cascade process down to the year scale. Then, on the next level of this cascade (i.e. a scale of six months) we need to "re-order" random multipliers in such way that among the two next random variables the largest one will be used first. So, the random multiplier creating the spring–summer singularity will be larger than the one that creates the autumn–winter singularity. It is important to note that such re-ordering will not modify the essential statistics, but breaks the time translation invariance cascade. If this re-ordering process is repeated for the next levels of the cascade process, then it will result in stronger periodicity. Once the cascade has been developed to a sufficiently small scale, one can upscale the resulting multifractal field up to the scale of interest, e.g. up to the month scale. Summation over many realizations of such multifractal fields could be interpreted as summing over the various basin contributions and corresponds to a simplified multifractal runoff model. Figure 11 illustrates a multifractal discharge yielded by this model. The "synthetic" multifractal discharge has both important properties: a visual seasonal periodicity and nonlinear DTM curves.

Fig. 9 Example of a log–log plot of the DTM as a function of the time scale ratio λ for $q = 1.5$ and $\eta = 1$.

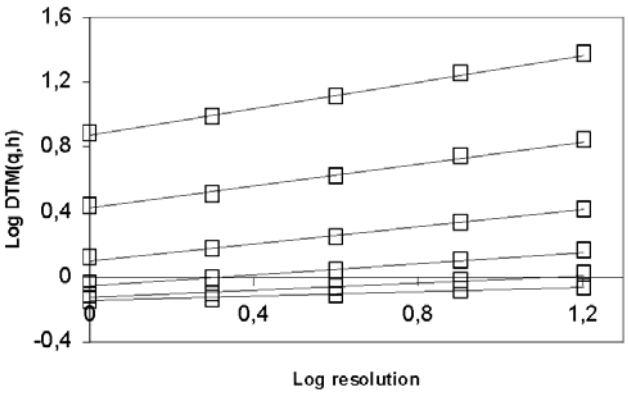


Fig. 10 A log–log plot of the DTM (average over 3321 gauges for year 1984–1985) as a function of the time scale ratio λ (from 1 month to 1 year) for $q = 1.5$ and different values of η. The straight lines demonstrate scale invariance for moments of monthly discharges and further lead to estimate of $\alpha = 1.56 \pm 0.3$ and $C1 = 0.28 \pm 0.08$.

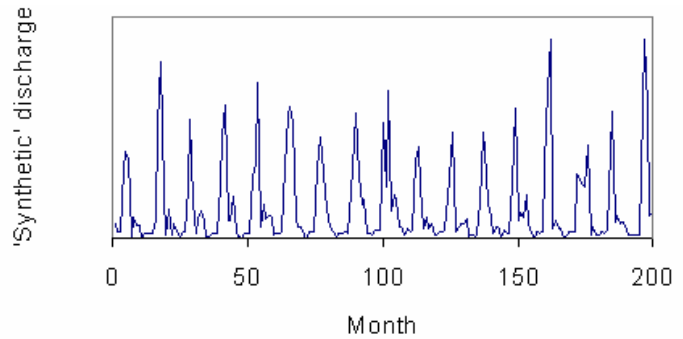


Fig. 11 Synthetic multifractal discharge time series exhibiting seasonal periodicity.

DISCUSSION AND POTENTIAL

The extreme variability over a wide range of scales suggested the potential of a multi-fractal analysis of river discharges in both the space and time domains, and of drainage areas and specific discharges in the space domain. Not only are river discharges and

drainage areas multifractal over the full range of scales but also they have the same index of multifractality: $\alpha = 1.35 \pm 0.2$. This value is comparable to $\alpha = 1.56 \pm 0.3$ that we obtained for the discharge time series analysis. Furthermore, we have demonstrated that a sum of the universal multifractal scaling function of drainage areas and of specific discharges may be well approximated by an empirical estimate of the scaling moment function of monthly discharges up to a rather high order of statistical moments. One important consequence of these results, which go well beyond the scope of earlier fractal drainage area studies, is presented in Fig. 12: the power-law of the probability distribution tail for the specific discharge distributions. This figure demonstrates that all this data have a unique parameter $qD = 6$ for the probability fall-off. It corresponds to an order of divergence of statistical moments in the universal multifractal framework (Hubert *et al.*, 2001).

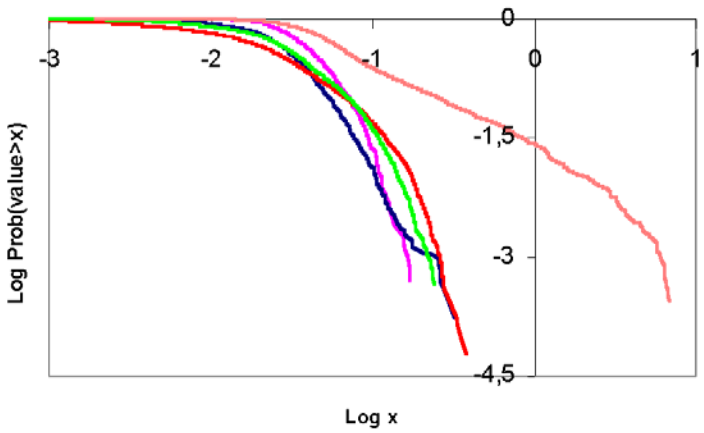


Fig. 12 Probability that monthly specific discharges exceed a given threshold. Log–log plot for space distributions in all the hydrological zones displayed in Fig. 1(b) (Anadyr and Kolyma; Lena; North European Russia and part of Scandinavia; Ob; Yenisei). For all these zones, note the tendency for the probability tail slope to get close to a value $qD = 6$.

We have argued that both the extreme variability over a wide range of space–time scales and seasonal periodicity are important features of river runoff. We have presented a generalization of the cascade processes in order to include them both: singularities are re-ordered in order such that the strongest ones occur during the season of strongest events. We showed that the corresponding model reproduces the main properties of the monthly discharge DTM curves thus giving a rationale to the latter. This is important for data analysis, since not only does it prevent misinterpretation of seasonal periodicity as a scale break, but also yields better estimates of the underlying universal multifractal exponents.

The absence of a scale break implies that the small-scale activity may introduce singular statistical behaviour in upscale observations, i.e. power-law fall-off of the probability distribution. The latter has highly significant consequences for the estimation of the natural variability, which is a preliminary requirement for any global change assessment, as well as for estimating hydrological risk for land-use planning or water structure design.

REFERENCES

Hubert, P., Bendjoudi, H., Schertzer, D. & Lovejoy, S. (2001) Multifractal taming of extreme hydrometeorological events. In: *The Extremes of the Extremes* (ed. by A. Snorrason, H. P. Finnsdottir & M. Moss), 51–56. IAHS Publ. 271. IAHS Press, Wallingford, UK.

Labat, D., Mangin, A. & Ababou, R. (2002) Rainfall–runoff relations for karstic springs: multifractal analyses. *J. Hydrol.* **256**, 176–195.

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

Monin, A. S. & Yaglom, A. M. (1975) *Statistical Fluid Mechanics*. MIT Press, Boston, Massachussetts, USA.

Pandey, G., Lovejoy, S. & Schertzer, D. (1998) Multifractal analysis including extremes of daily river flow series for basins one to a million square kilometres. *J. Hydrol.* **208**, 62–81.

Richardson, L. F. (1922) *Weather Prediction by Numerical Process*. Cambridge University Press, republished by Dover in 1965. Dover Press, Mineola, New York, USA.

Schertzer, D. & Lovejoy, S. (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* **D8**(8), 9693–9714.

Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y. & Marsan, D. (1997) Multifractal cascade dynamics and turbulent intermittency. *Fractals* **5**(3), 427–471.

Tchiguirinskaia, I., Hubert, P., Bendjoudi, H. & Schertzer, D. (2002) Multifractal modeling of river runoff and seasonal periodicity. In: *Preventing and Fighting Hydrological Disasters*. Timisoara, Romania.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D. & Pecknold, S. (1996) Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *J. Geophys. Res.* **31D**, 26427–26440.