Scaling, multifractals and predictions in ungauged basins: where we have been, where we are going?

D. SCHERTZER¹, P. HUBERT² & S. LOVEJOY³
1 CEREVE, Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée and Météo-France, Paris, France
daniel.schertzer@cereve-enpc.fr
2 CIG, Ecole des Mines de Paris, Fontainebleau, Paris, France
3 Physics Dept, McGill University, Montréal, Canada.

Abstract The multiplicity of scales, and hence the further development of scaling concepts and techniques, is at the core of the problem of Predictions in Ungauged Basins (PUB). Indeed, PUB can be restated in the following manner: given a partial knowledge of the input (atmospheric states, dynamics and fluxes) and of the media (basin) over a given range of scales, what can we predict for the output (streamflow and water quality) and over which range of scales? Therefore, we review the recent advances and challenges with respect to our concepts, analysis and modelling techniques of the scaling behaviour of the input, media and output. This helps us to identify a realistic strategy for investigating PUB.

Key words hydrology; prediction; ungauged basins; scales; scaling; multifractal

MULTIPLICITY OF SCALES AND PUB

A fundamental methodological problem

All agree that hydrological fields display an extreme variability over a wide range of nested space–time scales (e.g. Raudkivi, 1979; Tchiguirinskaia et al., 2004) and the scale ratio can easily reach $10^9$ (e.g. 1000 km–1 mm). This is particularly the case for the rain field, where drop distribution is inhomogeneous down to submetric scales (Lovejoy & Schertzer, 1990a; Desaulnier-Soucy et al., 2001; Lilley et al., 2002), whereas the external of cloud fields is of the order of planetary scale (Lovejoy et al., 2001; Lovejoy & Schertzer, 2006).

This is in sharp contrast with the inability of the classical methods to deal directly with such a wide range of scales. The latter are therefore compelled to introduce scale truncations and ad hoc parameterizations. Let us emphasize that these limitations are still relevant for the most advanced numerical simulation projects, such as the “Earth simulator” with its 1 km spatial resolution. In fact, the strong intrinsic limits of reductionist approaches in hydrology have often been pointed out (Beven, 1995; Wood, 1998).

These artificial sidesteps, nevertheless, lead to complex numerical codes that are both extremely difficult to transfer from one basin to another and to test with the help of empirical data that are not at the same scale. Furthermore, it is practically impossible to find an objective way to tune up the numerous parameters of these codes (e.g. Gosset & Gaume, 2002). As a consequence, predictions are reduced to fit and extrapolate past streamflow observations.

It is therefore not surprising that scaling approaches to hydrology have received a great deal of impetus. The general idea is that the search for invariance properties
across scales as fundamental hidden orders in hydrological phenomena should guide the development of data analysis and modelling methods (National Research Council, 1991). In other words, multiscale variability is no longer considered as a difficulty to be avoided at all cost, but as a consequence of a symmetry which should be uncovered so as to cast order in an apparent disorder. This is rather incompatible with the classical approaches: this symmetry is obviously artificially broken by any scale truncation.

**A restatement of PUB**

In our opinion, PUB not only provides an unprecedented opportunity to test the relevance and applicability of existing scaling concepts and techniques, but also requires their further developments.

There should be no ambiguity about the fact that PUB never had the goal of replacing the data missing due to the recent and drastic decline of hydrological *in situ* networks (Shiklomanov *et al.*, 2002; Vörösmarty *et al.*, 2002). These data are in any case indispensable and their (hopefully) temporary loss highlights the fact that these data should be much better and more intensively exploited. Indeed, too often they have been used for little more than tuning model parameters.

More fundamentally, PUB can be restated in the following manner: given a partial knowledge on the input (atmospheric states, dynamics and fluxes) and of the media (basin) over a given range of scales, what can we predict for the output (steamflow and water quality) and over which range of scales?

**Why multifractals?**

Let us briefly recall that hydrology has very greatly stimulated scaling ideas (for reviews see Schertzer & Lovejoy, 1991; Lovejoy & Schertzer, 1995; Rodriguez-Iturbe & Rinaldo, 1997; Sposito, 1998; Schertzer *et al.*, 2002). Indeed, from the 1950s onward particular attention was paid to scaling laws in hydrology (Hurst, 1951; Miller & Miller, 1955a,b; Hack, 1957). Scaling notions have been profoundly rejuvenated with the help of fractal concepts and models (Mandelbrot & Wallis, 1968, 1969; Lovejoy & Mandelbrot, 1985; Lovejoy & Schertzer, 1985) and much further with the help of multifractal concepts and multiplicative models (Schertzer & Lovejoy, 1987; Gupta & Waymire, 1993).

It is indeed symptomatic that scale dependence is rather ubiquitous in hydrology: clouds (Lovejoy, 1982), basins and rivers (Ijjasz-Vasquez *et al.*, 1994) are too tortuous to have a scale-independent area or perimeter. A similar scale dependence occurs for cloud cover (Gabriel *et al.*, 1988), precipitation occurrences (Hubert & Carbonnel, 1988, 1989), etc.

However, scaling of multifractal fields, e.g. the rain rate that is the fundamental quantity of interest for precipitation, is required to well beyond these geometrical observations. In this respect, multiplicative cascades have been particularly useful. Due to the fact that the cascade process is developed down to an infinitesimal scale, its limits are no longer a function but a (mathematical) measure (Halsey *et al.*, 1986). It is already the case for the simplest cascade model, the so-called β-model (Novikov & Stewart,
D. Schertzer et al.

1964; Mandelbrot, 1974; Frisch et al., 1978): scale by scale, the field clusters on a fractal set. Its limit is no longer a pointwise function, but a broad generalization of the Dirac impulse “function”, which is in fact a measure that concentrates a field in a point.

However, as soon as the geometrical and binary (active/not active structure) framework of the β-model is dropped out, a full hierarchy of activity levels is obtained, each of them clustered on a fractal set (Schertzer & Lovejoy, 1984). A scaling field can be understood as resulting from a hierarchy of fractal sets, hence the name multifractal (Benzi et al., 1984; Parisi & Frisch, 1985).

Multifractals can also be understood as a broad extension of geostatistics (Matheron, 1970; Delhomme, 1979): multifractals deal with random measures, instead of random functions, and the randomness is strongly non-Gaussian.

Examples of multifractal prediction: universality and self-organized criticality

To characterize a field as multifractal can be extremely useful since it may help to predict some of its fundamental features. For instance, there exist under fairly general conditions, “universal multifractals” that are attractive and stable limits of nonlinearly interacting identical multifractal processes (Schertzer & Lovejoy, 1987, 1997). Most of the physical fields of interest should be stable under these conditions, and therefore should be universal and characterized by very few universal exponents. The latter are physically significant, e.g. $C_1(C_1 \geq 0)$ characterizes the mean sparseness: it is the (fractal) co-dimension of the mean field (i.e. $C_1 = d - D_1$, where $D_1$ is the corresponding fractal dimension, $d$ the embedding dimension,) and an homogeneous field has $C_1 = 0$ ($D_1 = d$); $(0 \leq \alpha \leq 2)$ characterizes the degree of multifractality of the field: a mono/unifractal field has $\alpha = 0$, whereas the misnamed “lognormal” field has $\alpha = 2$, and in a general manner $\alpha$ measures how rapidly the sparseness increases as the level of activity increases.

Another general prediction can be drawn from the observed multifractal behaviour of a field. Indeed, Schertzer & Lovejoy (1992) showed that the tail of the cumulative probability distribution should be a power-law rather than an exponential law. Its exponent $q_D$ corresponds to the critical order of statistical moments, i.e. the mathematical expectation of any $q$th power of the field will be infinite for any $q \geq q_D$, although its empirical estimate on a finite sample will be finite and spurious (for more discussion see Schertzer et al., 2002). This singular statistical behaviour (Schertzer et al., 1993a,b) can be discussed in relation to the notion of self-organized criticality (Bak et al., 1987, 1988; Bak & Tang, 1989; Bak & Chen, 1991), which is also based on scaling.

WHAT DO WE KNOW ABOUT SPACE–TIME VARIABILITY?

Input variability

Multifractal behaviour of the rainfield was analysed on various precipitation time series ranging from milliseconds to centuries, as well on a few spatial series ranging from metres to planetary scales (Schertzer & Lovejoy, 1987; Gupta & Waymire, 1990, 1993; Ladoy et al., 1991, 1993; Fraedrich & Larnder, 1993; Hubert et al., 1993, 1995, 2001; Tessier et al., 1993; Olsson, 1995, 1996; Carsteaunu & Foufoula-Georgiou, 1996;
Scaling, multifractals and predictions in ungauged basins

Harris et al., 1996; Olsson & Niemczynowicz, 1996; de Lima, 1998; Bendjoudi et al., 1997; Schmitt et al., 1998; Biaou, 2002; de Lima et al., 2002; Mouhous, 2002). This effort also bears on the understanding of the radar and satellite measures (Lovejoy & Schertzer, 1990b; Lovejoy et al., 1996; Féral & Sauvageot, 2002) as well as the difficulties faced by classical calibration methods (Gabriel et al., 1988; Giraud et al., 1986; Tessier et al., 1994).

Particular attention was paid to determine the universal exponents $C_{1}$ and $\alpha$ of the rainfall rate. In spite of the development of rather robust statistical methods (Lavallée et al., 1992, 1993), their determination remains difficult due to the extreme and non-classical variability that they characterize. Furthermore, there is the additional and important problem of the zeroes. Empirically, the zeroes greatly affect the low rain rate statistics, hence methods for estimating $\alpha$ will be sensitive to the way the measuring instrument handles the problem; hence estimates of $\alpha$ range from 0.5 to 0.75 for time series and from 0.9 to 1.5 for space. There is the need to pursue this empirical investigation in order to reduce the scatter of the present results, in particular for space and time comparison, although the latest results from the TRMM reflectivities (Lovejoy & Schertzer, 2007) indicate that a single multiplicative process coupled with low instrumental detection thresholds can accurately account for the reflectivity (and hence presumably precipitation) statistics down to at least 5 km.

Nevertheless, independently of more precise estimates of the universal exponents, an exponent $a_{qD} \approx 3$ for the exceedence probability tail has been shown to be rather universal for the rain-rates time series (Hubert et al., 2001). As a consequence, the asymptotic law of the extremes is of Fréchet type rather than Gumbel type (Hallegatte et al., 2002; Schertzer et al., 2006; 2007).

Note that we are rather at the beginning of the multifractal analysis of space–time variability (Marsan et al., 1996; Over & Gupta, 1996; Biaou, 2002), although the corresponding methods have existed for a while (e.g. space–time generalized scale, Schertzer & Lovejoy, 1985). Furthermore, it will be important to characterize the scaling inter-relations between rain and other atmospheric fields (dynamics, clouds, moisture fluxes, etc.).

Output variability

Similar investigations have been progressively extended to flow rates (Turcotte & Greene, 1993; Tessier et al., 1996; Pandey et al., 1998; Hubert et al., 2002; Labat et al., 2002; Tchiguirinskaia et al., 2002) and much remains to be done for water quality. Universal exponents and critical exponents of moment divergence were estimated, but they exhibit a larger scatter than their rainfield counterparts. This scatter is presumably physically due to the variability of basins, as well as the necessity to perform an adequate renormalization of the flow rates with respect to the basin size (Tchiguirinskaia et al., 2002, 2007).

Basin variability

The spatial variability of basins has mostly been investigated in the framework of fractal geometry, in particular for their geomorphology and river networks (Shreve,
1966, 1969; LaBarbera & Rosso, 1987; Tartabon, 1988; Robert & Roy, 1990; Tartabon et al., 1991; Ijjasz-Vasquez et al., 1993; Maritan et al., 1996). Nevertheless, Klinkenberg & Goodchild, 1992; Lavallée et al., 1993; Verge & Souriau, 1994; Lovejoy et al., 1995; Pecknold et al., 1997; Gagnon et al., 2006) performed multifractal analyses of the topography. The latter is known to be of prime importance for the basin response (Beven & Kirkby, 1979; Tartabon et al., 1991), in particular for wetlands (Tchiguirinskaia et al., 2000). It turns out that topography multifractality seems to be quite universal from at least 40 m up to planetary scales and somewhat surprisingly its multifractality is rather extreme ($\alpha \approx 1.8$), whereas its mean fractality is rather low ($C_1 \approx 0.12$).

However, other soil properties have to be taken into account. This in particular is the case of hydraulic conductivity whose fractal properties (Wheatcraft & Tyler, 1988; Tyler, 1990), then multifractal properties (Tchiguirinskaia, 2002) have been analysed and show much more universality than previously believed.

WHERE DO WE GO?

Let us again emphasize that further empirical analyses are needed for empirically estimating a few fundamental scaling exponents, and therefore data. On the other hand, and rather in parallel, it is important to further develop stochastic models in order to better understand—with the help of either their analytical properties or their numerical simulations—the interrelations between various fields, in particular for their extremes, how to up/downscale or how to condition the large scales, the meaning of remotely sensed measurements, the predictability limits, and how to proceed to stochastic forecasts (Schertzer & Lovejoy, 2004).

Multifractal modelling

The basic numerical algorithm for making continuous (in scale) multifractal models, including non-conservative fields, was first described in Schertzer & Lovejoy (1987), extensions to downscaling in Wilson et al. (1991), to linear generalized scale invariance in Pecknold et al., (1993), to causal space–time modelling in Marsan et al., (1996) and Schertzer et al. (1997), as well as extensions to non scalar fields in Schertzer & Lovejoy (1995). We expect more or less straightforward developments in order to obtain more and more adequate modelling of the hydro-meteorological input that should be used for effective multifractal forecasts, e.g. with the help of incomplete radar and/or satellite data.

The key developments will be at the level of basin response modelling. Whereas a (linear) fractional integration of rain-rates yield some realistic-like simulations of river runoff in certain cases (Tessier et al., 1996; Pandey et al., 1998), more careful analysis shows that it requires other developments. For instance, far beyond the scope of fractal generation of river networks (Scheidegger, 1967; Takayasu et al., 1988), one has to take into account the multifractality of the drainage area (Tchiguirinskaia et al., 2002), which up to now has been has understood as a mere consequence of a finite-size effect (Maritan et al., 1996).
Anomalous transport in complex media

On the other hand, the associated anomalous transport properties could be approached along the lines of a generalized diffusion in complex media in order to explain the fractal behaviour (Kirchner et al., 2000), in fact the presumably multifractal behaviour of contaminant concentration fluctuations and more generally of water quality.

The first level of complexity is obtained with the help of a classical disorder, i.e. Gaussian, but whose intensity is extremely inhomogeneous in space. One then may obtain a classical transport equation (advection–diffusion, Fokker Planck (Van Kampen, 1981), but with extremely variable coefficients (Machta, 1981; Zwanzig, 1982; Havlin & Ben Avraham, 1987; Kavvas & Karakas, 1996). The extreme case corresponds to a multifractal intensity of the microscopic disorder (Meakin, 1987; Marguerit et al., 1997; Lovejoy et al., 1998), which indeed yields anomalous diffusion laws. This has been proposed for river modelling (Meakin et al., 1991), although without any concrete application.

Another level of complexity is obtained by considering a strongly non-Gaussian microscopic disorder, more precisely a Levy white-noise. Fluctuations are so important, that the “microscopic” refers only to the particle size, not to its effect. One obtains a Fractional Fokker Planck equation, in fact a fractional diffusion–advection equation (Zaslavsky, 1994; Chechkin, 1995; Compte, 1996; Yanovsky, 1997); in particular, the classical Laplacian diffusion operator is raised to a fractional power. Benson et al. (2000, 2001) strongly argued for its applicability to subsurface transport.

Finally, considering a Lévy disorder with an inhomogeneous intensity can combine both complexity levels. One obtains (Schertzer et al., 2001) an inhomogeneous fractional Fokker-Planck equation, whose properties have been not yet fully explored.

An investigation strategy for PUB

The discussion above defines in fact a dual strategy to investigate PUB. On the one hand, we have to analyse larger and large databases (including remote sensed data) to better characterize the multiscale variability of the inputs, the basin and the flow output. This is indispensable to better assess what are the common features (e.g. universal exponents) and the differences. In parallel and in close interaction with these empirical developments, we need to further develop and test our modelling capacities. It is also important to foresee somewhat autonomous modelling developments to define conceptual PUB problems that will be tractable enough to give deep insights into the more involved and realistic PUB problems.

REFERENCES


Scaling, multifractals and predictions in ungauged basins


