Theory and Measurements of the Distribution of Rainfall in Space and Time

A.W. Seed
CRCCH, Bureau of Meteorology, GPO Box 1289K, Melbourne, 3001, Australia (a.seed@bom.gov.au)

Abstract: The distribution of rainfall in space and time is well known to be both variable and scale dependent. Models based on scale invariance have been shown to provide parsimonious descriptors of the distribution of rainfall over a wide range of scales in both time and space. This paper will present some of the methods that are used to describe the scaling behaviour of rainfall and examples of the analysis of Australian radar rainfall data will be presented to demonstrate the scaling nature of rainfall. The scaling nature of rainfall affects both the ability to forecast rainfall and the selection of stochastic methods to generate plausible fields of rainfall. Examples of methods that exploit the space and time scaling of rainfall to produce short duration rainfall forecasts and stochastic simulations will be presented.

Keywords: Rainfall variability; Weather radar, Scaling

1. INTRODUCTION

Rainfall is the result of complex atmospheric phenomena, and therefore is itself complex displaying both variability and intermittency over a wide range of scales. The statistical properties of rainfall in space and time are highly dependent on the scale of the accumulation or average, making it an "intriguing and challenging process to study" [Foufoula-Georgiou, 1996]. In general, there are two generic types of rainfall models viz. those based on clustered point process models using the notion of rain cells [Cox and Isham, 1996] and those based on treating the rain field as scaling or multifractal process. Scaling models of rainfall provide "attractive and parsimonious" [Foufoula-Georgiou, 1996] representations that are able to generate rain fields that are able to reproduce the observed relationship between the statistics of precipitation amounts at different spatial and temporal scales. While some argue that rain fields (or more accurately the fluctuations of rain) are multifractal in a fundamental sense e.g. Lovejoy and Schertzer [1995], it is sufficient to note that they are able to reproduce a reasonable approximation of key observed statistical characteristics over the range of scales usually found in hydrology, and are therefore useful in modelling and simulating rain fields for hydrological applications. This paper will start with a very short outline of the theory that underpins the analytical methods but will include useful references to more detailed treatments. Thereafter, the use of scaling models to characterise the statistical nature of rainfall, to generate plausible random fields, and to produce conditional simulations will be discussed and conclusions will be drawn. This paper is not an encyclopaedic treatment of the subject, but rather an introduction to scaling analysis and the presentation of some models that the author has found useful.

2. BASIC THEORY

A random field will display scaling characteristics if and only if the power spectrum $P(f)$ obeys the power law

$$P(f) \propto f^{-\beta} \qquad (1)$$

The scaling can be characterised by spectra of exponents, the spectra for the scaling of the moments, $K(p)$, defined as

$$\left\langle F_l^q(x)\right\rangle \propto l^{-K(q)} \qquad (2)$$

where $F_l^q(x)$ denotes the average of the field over the interval l centred at x, and < > denotes ensemble average, and the spectra for the scaling of the generalised structure function, $\zeta(q)$, defined as

$$\left\langle \left|F(x+l)-F(x)\right|^q \right\rangle \propto l^{\zeta(q)} \qquad (3)$$

Scaling can also be expressed in terms of the scaling of the probability distributions for the field averaged over scales l

$$\Pr(F_l \geq l^{\gamma}) \approx l^{-c(\gamma)} \qquad (4)$$

where $c(\gamma)$ is the spectrum of scaling exponents for the probability distribution, and ≈ represents equality to within slowly varying functions of l

[Lovejoy and Schertzer, 1995]. The exceedence probabilities of multifractals have asymptotically hyperbolic tails so that $\Pr(R > r) \propto r^{-q_{cr}}$ [Mandelbrot, 1974]. Here q_{cr} is a critical exponent indicating the moment beyond which moments are divergent and $K(q)$ becomes a straight line [Harris et al., 1996].

It can be shown fields are multiscaling as defined by equation (2) if the slope of the power spectrum, β, is less than D, the dimension of the support ($D = 1$ for time series, 2 for spatial fields). In the case for $\beta > D$, the scaling is characterised through the generalised structure function defined by equation (3) and is said to be multiaffine. When a field with $\beta > D$ undergoes a small scale absolute gradient transformation the resulting ΔF field will be multiscaling with a power spectrum $\beta < D$.

Estimation of $K(q)$ simply involves finding the q^{th} moment of the field at the smallest scale and then averaging the field over successively larger scales. The $K(q)$ curve is only defined over a finite range of uniform scaling and care must be taken to verify that the field does in fact scale over a useful range. The analytical form for $K(q)$ can be derived theoretically for specific cascade constructions e.g. Universal Multifractals of Lovejoy and Schertzer [1995], or the Bounded Lognormal Cascades of Menabde [1998]. The advantage in these approaches is that they provide a convenient framework for the characterisation of highly non-linear fields over a wide range of scales by means of a small number of parameters. For example, C_1, the derivative of $K(q)$ at $q = 1$, is a measure of the intermittency of the field (the codimension of the set of points that are greater than the mean). If C_1 is close to zero the field behaves as a field of white or $1/f$ noise which fills the embedding space, whereas if C_1 is close to D then the field is intermittent and the flux is concentrated on a small fraction of the area.

A major problem with verification of a multifractal or multiaffine behaviour in rainfall data sets using equations (2) or (3) is that the analysis requires long sequences of statistically homogeneous rainfall data. Long sequences of rainfall data may not be scaling over the entire period and usually include sequences from several different rainfall processes, each with significantly different statistical and scaling properties. Fabry [1996] found that a time series of rainfall intensities from the Montreal area did not show a scaling power spectrum at scales that were greater than 5 days, and therefore scaling models are essentially models for the disaggregation of rainfall events. Rainfall records are also prone to measurement errors and instrumentation artefacts, and great care needs to be taken to eliminate as much corrupt data as possible prior to analysis. This is particularly true when using radar rainfall data to investigate the scaling of the moments or the probability distribution. A review of the impact of measurement noise and non-stationarity on scaling techniques is given by Harris et al. [1998].

Most of the theory and models for scaling invariance are based on the random multiplicative cascades, which arose in the theory of turbulence. The random multiplicative cascade can be constructed as a discrete cascade e.g. Over and Gupta [1996], Menabde [1998], or as a continuous cascade e.g. the Universal Multifractals of Lovejoy and Schertzer [1995]. An alternative approach is to use the wavelet decomposition e.g. Perica and Foufoula-Georgiou [1996]. Substantial reviews of the models and related theory can be found in Lovejoy and Schertzer [1995] and Foufoula-Georgiou [1996].

A discrete random cascade is constructed as follows: At an arbitrary n^{th} step in the process, a square of size $l_n = L2^{-n}$ with mean areal rainfall F_n is divided into four equal squares, each with size $l_n/2$ and value $F_{n+1} = W(i)F_n$ where the $W(i)$ are random numbers drawn from some generator. If the weights are selected subject to the constraint that $\sum_{i=1}^{4} W(i) = 1$, then the cascade is said to be microcanonical and the resulting field will have simple scaling characteristics.

A less restrictive condition is to set the mean of the generator equal to 1, the cascade is said to be canonical and resulting field will be multiscaling if the weights are independent and identically distributed for all levels of the cascade, or multiaffine if the width of the generator is reduced with scale according to a power law [Menabde, 1998]. A schematic of the construction of a self-similar discrete multiplicative cascade taken from Lovejoy and Schertzer [1995] is shown in Figure 1.

3. MULTIFRACTAL CHARACTERISATION OF RAIN FIELDS

Harris et al. [1996] used β, C_1, and q_{cr} to characterise the statistical nature of precipitation as a function of location in a mountain range. A network of rain gauges with 15-second resolution was installed on the west coast of the South Island of New Zealand. The analysis showed a systematic trend of decreasing intermittency and extreme values with increasing altitude (or proximity to the crest of the mountain). For example, β decreased from a value of 1.5 at the coast to 0.95 at the main divide pointing to a relative reduction in the power of low frequency structures at the main divide. C_1, the measure of intermittency, decreased from the coast to the main divide as did q_{cr}, a measure of the width of the distribution. Nagata [2001] used β and C_1 to classify fields of hourly rainfall accumulations into "widespread" and "convective" rainfall classes. Characterisations of this type are interesting since they quantify the statistical differences that exist between rain fields that have different meteorological origins.

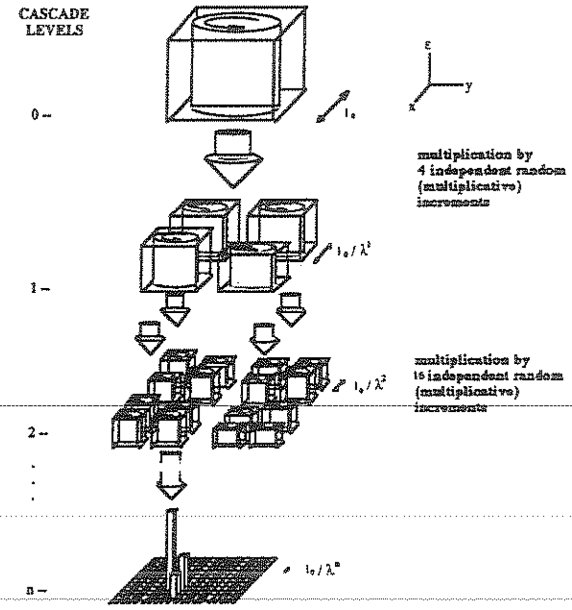

Figure 1. A schematic of the construction of a discrete multiplicative cascade, taken from Lovejoy and Schertzer [1995].

4. INTENSITY-DURATION-FREQUENCY RELATIONSHIPS

The relationship between rainfall intensity, duration, and frequency (IDF) has been of considerable interest to engineers for over a century, and a large number of empirical relationships have been proposed over the years. The scaling hypothesis offers a consistent framework for an analysis of the IDF characteristics for extreme rainfall. Burlando and Rosso [1996] were pioneers in applying a scaling model to an annual maximum series of rainfall depths over a range of accumulations. Menabde et al. [1999] applied the simple scaling hypothesis to model the cumulative distribution function for the annual maximum series of rainfall intensity. Assuming the EV1 distribution for the annual maximum series, the IDF relation was found to be

$$i = \left[\mu - \sigma \log\{-\log(1 - 1/T)\}\right] / d^{\eta} \qquad (5)$$

where i is the mean rainfall intensity with duration d hours and annual recurrence interval T years.

The μ and σ parameters of the EV1 distribution fitted to the annual maximum series for rainfall durations of 30 minutes up to 24 hours were found to follow a simple power law $(d/D)^{\eta}$, where η depends only on the climate. It is possible to use a long sequence of daily data to calculate μ and σ, and a shorter sequence of high time resolution data in the climatic region to estimate η since the intensities of events with the same annual recurrence interval but different durations will scale with the same exponent η.

5. MODELLING RAINFALL IN SPACE AND TIME

Theories of space-time rainfall based on scaling ideas have only recently emerged (e.g. Bell [1987], Over and Gupta [1996], Marsan et al. [1996], Seed et al. [1999], Pegram and Clothier [2001]). A general feature of atmospheric turbulence is that for a feature with a given size l, there is a typical lifetime or correlation length $\tau_l \propto l^{1-H}$ [Marsan et al., 1996], leading to anisotropic scaling in space and time. This space-time model is quite different to the usual meteorological phenomenology, which assumes that there is a hierarchy of qualitatively different dynamical mechanisms that depend on scale [Schertzer et al., 1997]. This complicates space-time models since the scaling in space is different from that in time and this anisotropy is difficult to quantify since it must be estimated in Lagrangian rather than Eulerian coordinates.

A further complication arises from the fact that cascade models are essentially event-based models since rainfall typically does not scale convincingly beyond a couple of days. The Pegram and Clothier [2001] model first models the arrival and duration of the events as an alternating renewal process and then disaggregates each event into a sequence of rain fields. The disaggregation is done by using an anisotropic power law spectral filter on a cube of Gaussian noise to achieve the observed correlations in space and time, and then exponentiating to obtain a time series of correlated lognormal fields. The time series of mean areal rainfall over the entire field is modelled by first generating a time series of the wetted area ratio based on an AR(6) model, and then using a derived relationship between the mean and the wetted area ratio to estimate the mean and variance of the field at a particular time step. Each field in the time series of the event is then renormalized so as to produce the correct field mean and variance.

The Seed et al. [1999] model "Motivate" is an event model that is intended to be used to generate plausible realisations of a design storm. The model consists of three components viz. a broken line model [Seed et al. 2000] to generate the time series of mean areal rainfall, a model based on a bounded log-normal cascade to generate the spatial pattern, and an AR(2) model for the temporal evolution of the cascade weights.

The model has been used to generate plausible 24-hour 1 in 5 year ARI design storms for the Melbourne area [Seed et al. 2001a]. Melbourne Water is required by the EPA to progressively upgrade their sewerage system so that is able to contain the seepage into the system resulting from a 1 in 5 year ARI storm. At present the design storm hyetograph is generated using the standard Australian Rainfall and Runoff (ARR) method of first estimating the 1 in 5 year storm total using rain gauge data, applying an areal reduction factor to convert the point measurement into an equivalent areal measurement over the 1600 km^2 catchment, and then distributing the storm total in 5-minute time steps using a standard temporal pattern. An infiltration model is then used to predicted flows at critical locations for the design storm.

The major problems with this approach are firstly that the rainfall pattern is uniform over the entire network leading to very high estimates of pipe flow, and secondly only one design storm scenario can be generated. The advantage of using a space-time model to generate an ensemble of design storms, each with the same mean and duration, is that they can be used to evaluate the variability in the hydrological response that arises from the differences in the small scale details in the rain field.

Motivate requires the gross statistics for the event (duration, mean intensity, variance at the small and large scale, and advection velocity) and scaling descriptors that control the space and time correlation structure of the field. The mean intensity for a given duration of the design storm was estimated using rain gauge records. The scaling parameters, variance of the rain field, and advection velocity were estimated using radar data.

The rain gauge record was used to identify events that had an ARI that was close to 1 in 5 years. An analysis of the synoptic situation for each event was undertaken and the storms were classified into 6 classes. The radar data archive was searched and significant storms for three of the event classes were identified. The radar data were used to estimate the speed and direction of the field advection as well as the scaling exponents. Two of the events could be characterised as widespread rainfall associated with rain bands ahead (in a NW airstream) or behind a cold front (in a SW airstream), while the third was convective rainfall embedded in a NE airstream. The scaling exponents were quite similar for the first two cases, while rainfall associated with the third event was more variable and intermittent.

6. PREDICTABILITY AND FORECASTING RAIN FIELDS

The characterisation of a rain field as an ensemble of structures of scales between the single pixel and the outer scale of the field, characterised by dynamic scaling where the life time of a structure is a power law of it's scale, leads to an intuitive picture of the predictability of a rain field as a function of scale [Marsan 1996]. Since there is no information in the current image about a structure after a period that is longer than it's lifetime, the optimal strategy is to smooth structures once their lifetime has been exceeded.

Dynamic scaling implies that one needs to know the field on increasingly large scales in order to predict further into the future. Figure 2 shows an example of the lifetime (correlation length) of a structure as a function of scale for two tropical cyclones and an analysis of two cases of high-resolution (100 m, 6 s) radar data from the Physics Department, Auckland University, mobile radar.

These ideas have been exploited by Seed [2001b] to develop the Spectral Prognosis (S_PROG) model. Rainfall is related to radar reflectivity through a power law, so a multiplicative cascade of rainfall can be transformed into an additive cascade of radar reflectivity in dBZ. The cascade of random fields is calculated from a measured field by using notch filters in the frequency domain to disaggregate the ensemble of scales that are present in the field into a hierarchy of levels in a cascade that will then sum back into the observed field. The mean advection of the field is calculated by finding the displacement between successive images that maximises the correlation between the images.

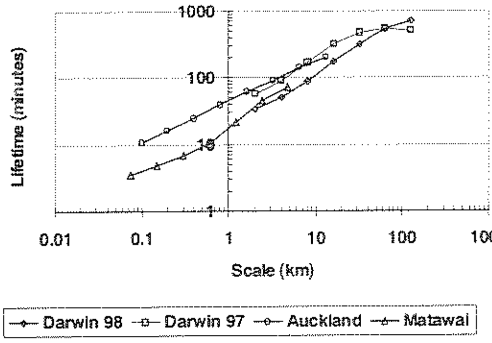

Figure 2. Dynamic scaling for two tropical cyclones (Darwin 97, Darwin 98) and two cases using high-resolution radar data from New Zealand (Auckland, Matawai).

The temporal development of each level in the cascade is modelled using an AR(2) model for each level in the cascade. The parameters for the hierarchy of AR(2) models are calculated at each time step by using the most recent estimates of the Lagrangian lag 1 and 2 auto correlations to solve the Yule-Walker equations, applying heuristic rules to maintain stationarity. This has the effect of making the model quite adaptive and generic, learning the characteristics of each rainfall event as it unfolds.

The forecast fields are then generated by using the hierarchy of AR(2) models (with out the noise if a single forecast is required, or with the noise term if conditional simulations are required) to predict the future state of the cascade, and advecting the field by the measured displacement. The forecast fields are then renormalised so as to maintain the observed conditional mean rain rate and raining area, and are converted from radar reflectivity into rainfall intensity in the usual manner and with the usual problems. Examples of the model output are shown in Figures 3 and 4.

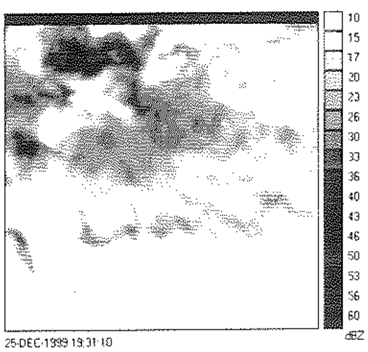

Figure 3. 10-minute forecast using S_PROG

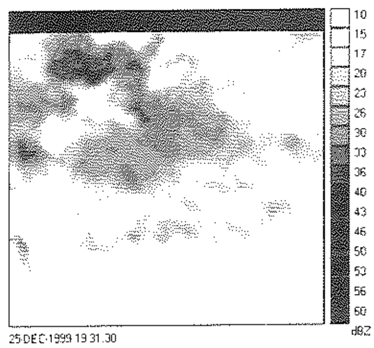

Figure 4. 30-minute forecast using S_PROG

7. FUTURE ISSUES AND PROBLEMS

All current analysis methods require very large samples of data. Problems arise therefore from non-stationarity in the data and contamination with measurement noise. The lack of robust methods to evaluate either the scaling of the moments or the generalised structure function for rainfall is a major problem that requires further research.

The moment scaling methods of double trace moments by Lovejoy and Schertzer [1995] are attractive in that they rely on successive averaging over larger spatial scales, thereby reducing the impact of small-scale noise, but have unsolved issues regarding the treatment of the zero values. The generalised structure function is difficult to estimate, particularly for high moments, and this sampling variability needs to be understood in order to assess the significance of the difference between two structure functions. Very little work has been done to exploit the $c(\gamma)$ spectra of exponents for the scaling of the probability distributions.

All current models assume that the scaling is homogeneous in both space and time during an event. This is most unlikely to be the case but very little work has been done on either quantifying the variability of the scaling descriptors in space and time, or relating them to meteorology, and no work has been published on models that can generate spatially inhomogeneous scaling fields of rainfall. Models based on wavelets are very attractive in this regard as the wavelets, unlike the Fourier methods, are located in space and therefore could be more easily adapted to spatially non-homogeneous scaling models.

The spatial organisation of rainfall accumulations on the ground is dependent on both the nature of the instantaneous rain field and the velocity with which the field is advected over the area. To some extent, the accuracy of small-scale precipitation forecasts is limited by the predictability of the advection field at small-scales. The variability of field advection both during an event and from one event to the next has yet to receive adequate attention, and the predictability of the advection vector as a function of scale has yet to receive any attention.

Scaling models of rainfall are able to provide powerful representations of rain fields in both space and time. These models are generally parsimonious and robust, capturing many of the features in rainfall that are important to hydrology. The theory is still very young and therefore a standardized argot has yet to emerge making it difficult to relate the research from one group to that from another. While the jury is still out on the question of whether rain fields are multifractal in a fundamental sense, multifractal models are still useful representations of the measurements and worthy of further development.

8. References

Bell, T.L., A space-time stochastic model of rainfall for remote sensing studies, *J. Geophys. Res.*, 92(D8), 9631-9643, 1987.

Burlando, P, and R. Rosso, Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation, *J. Hydrol.*, 187, 45-65, 1996.

Cox D.R., and V. Isham, Stochastic spatial-temporal models of rain, In O.E. Barndorff-Nielsen, V.K. Gupta, V. Pérez-Abreu, and E. Waymire, *Stochastic Methods in Hydrology: Rain, Landforms and Floods*, World Scientific, Singapore, 1996.

Fabry, F., On the determination of scale ranges for precipitation fields, *J. Geophys. Res.*, 101(D8), 12819-12826, 1996.

Foufoula-Georgiou E., On scaling theories of space-time rainfall: some recent results and open problems, In O.E. Barndorff-Nielsen, V.K. Gupta, V. Pérez-Abreu, and E. Waymire, *Stochastic Methods in Hydrology: Rain, Landforms and Floods*, World Scientific, Singapore, 1996.

Harris, D., M. Menabde, A.W. Seed, and G.L. Austin, Multifractal characterisation of rain fields with a strong orographic influence, *J. Geophys. Res.*, 101(D21), 1996.

Harris, D., A.W. Seed, M. Menabde, and G.L. Austin, Factors affecting multiscaling analysis of rainfall time series, *Nonlinear Processes in Geophys.*, 4(3), 137-156, 1998.

Lovejoy S., and D. Schertzer, Multifractals and rain, In Z.W. Kundzewicz, *New Uncertainty Concepts in Hydrology and Water Resources*, Cambridge University Press, 1995.

Mandelbrot, B., Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier, *J. Fluid Mech.*, 62(2), 331-358, 1974.

Marsan, D., D. Schertzer, and S. Lovejoy, Causal space-time multifractal processes: Predictability and forecasting of rain fields, *J. Geophys. Res.*, 101(D21), 26333-26346, 1996.

Menabde, M., Bounded lognormal cascades as quasi-multiaffine random fields, *Nonlinear Processes Geophys.*, 5, 63-68, 1998.

Menabde, M., A.W. Seed, and G.G.S. Pegram, A simple scaling model for extreme rainfall, *Water Resour. Res.*, 35, 335-339, 1999.

Menabde, M., and M. Sivapalan, Modelling of rainfall time series and extremes using bounded random cascades and Levey-stable distributions, *Water Resour. Res.*, 36, 3293-3300, 2000.

Nagata, K., *Spatial Variability and Rainfall Estimation*, Unpublished MSc thesis, Physics Department, University of Auckland, 2001.

Over T.M., and V.K. Gupta, A space-time theory of mesoscale rainfall using random cascades, *J. Geophys. Res.*, 101(D21), 26319-26331, 1996.

Pegram, G.G.S., and A.N. Clothier, High resolution space-time modelling of rainfall: the "String of Beads" model, *J. Hydrol.*, 241, 26-41, 2001.

Perica, S., and Foufoula-Georgiou, E., Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *J. Geophys. Res.*, 101(D21), 26347-26361, 1996.

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

Seed, A.W., R. Srikanthan, and M. Menabde, A space and time model for design storm rainfall, *J. Geophys. Res.*, 104(D24), 31623-31630, 1999.

Seed, A.W., C. Draper, R. Srikanthan, and M. Menabde, A multiplicative broken-line model for time series of mean areal rainfall, *Water Resour. Res.*, 36, 8, 2395-2399, 2000.

Seed, A.W., B. Taylor, and U. Maheepala, *The simulation of 1 in 5 year ARI design storms for Melbourne*, Unpublished report to Sewerage Planning, Melbourne Water, April 2001a.

Seed, A.W., A dynamic and spatial scaling approach to advection forecasting, Submitted to *J. Appl. Meteorol.*, August 2001b.