Journal of Hydroscience and Hydraulic Engineering
Vol. 22, No.2 November, 2004, 59-78

ANALYSIS OF MULTIFRACTAL PROPERTIES OF TEMPORAL AND SPATIAL PRECIPITATION DATA IN JAPAN

By

Y. Kuzuha

Faculty of Bioresources, Mie University, 1515 Kamihama, Tsu, Mie 514-8507, Japan

T. M. Over

Department of Geology / Geography, Eastern Illinois University, 600 Lincoln Ave., Charleston, Illinois 61920, U.S.A.

K. Tomosugi

Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

and

T. Kishii

Division of Environmental Engineering, Kanazawa Institute of Technology, 7-1 Ohgigaoka-Nonoichi, Ishikawa 921-8501, Japan

SYNOPSIS

First, we examined scaling and multifractal properties of temporal precipitation data with an emphasis on the structure of the rainy vs. non-rainy periods, using the so-called β model. These data have scaling properties in a range of 1 day $\leq L_n \leq$ 32 days. We characterize the scaling of the rainy areas with the single parameter of the β model and classified surface observation stations into two groups according to the parameter; namely, along the Pacific Ocean and Japan Sea coasts. The Pacific Ocean data are more intermittent and peaky than the Japan Sea data and exhibit more multifractality than the latter.

Second, we analyzed the time variation of scaling properties of several months of spatial radar data using the intercept of the regression estimate of the scaling of the fractional wetted area. These results showed that over the sea in the summer, there was a tendency for a different scaling of small vs. large scales.

Finally, we extracted four independent precipitation events from the whole dataset and examined whether or not a one-to-one function can represent the relationship between mesoscale forcing and the parameter of the β model. We conclude that there is a one-to-one functional relationship if we consider each precipitation event separately and exclude multiple independent precipitation events.

INTRODUCTION

The U. S. Federal Emergency Management Agency (FEMA) map for Presidential Disaster Declarations (http://www.bakerprojects.com/fema/) indicates that, in the United States, there is a high ratio of number of floods and severe storms to the total number of Presidential Disaster Declarations. These natural disasters are also severe in many developing countries. Although some sort of disaster prevention specialists in Japan are interested only (or at least almost only) in earthquakes, it goes without saying that water-related disasters, which include severe storms, landslides and floods, also threaten our lives, economic foundation and social systems, and that ignorance of these natural disasters leads to loss of life and property. Flooding is the outcome of several hydrological processes. Spatial variability of annual maximum floods is influenced by spatial variability of rainfall in large basins and by the relationship between catchment response time and the duration of rainfall in small basins (Gupta and Dawdy (7), Robinson and Sivapalan (23) and Kuzuha *et al.* (unpublished manuscript)). Blöschl and Sivapalan (2) indicated that the variability of annual maximum floods is influenced by interplay among numerous hydrological processes, and we know that other water-related disasters are also influenced by numerous hydrological processes. However, heavy rainfall is the chief cause of occurrence of these disasters. Many meteorologists and hydrologists have made efforts to gain an understanding of the temporal and spatial distribution of rainfall. First, we will review the *random cascade*, a powerful tool for modeling the temporal and spatial variabilities of rainfall, referring to Gupta and Waymire (9) and Over and Gupta (22).

According to Gupta and Waymire (9), the following idealized geometric structure has guided the development of statistical models of space-time variability of rainfall for the past 20 years. A hierarchical spatial structure of rainfall intensity is widely observed in mesoscale rainfall: the smallest spatial structures consist of clusters of regions (cells) of high-intensity rainfall that are embedded within regions (small mesoscale areas) of low rainfall intensity that are in turn embedded within large mesoscale areas in which rainfall intensity is even lower. Austin and Houze (1) were the first to observe empirical evidence of this structure within storms of a variety of synoptic types occurring in New England. Attempts to model this hierarchical spatial structure are classified into two types of approach (Over and Gupta (21)). One is the physical approach, which involves the numerical solution of a set of equations describing the dynamics and thermodynamics of the atmosphere. The other is the statistical approach, which involves the use of stochastic methods to describe the hierarchical spatial organization of rainfall intensity at mesoscale level and which we will describe in this paper. Statistical methods may be further divided into three types: (a) point process models, beginning with LeCam (14); (b) random field models, such as the Gaussian model of Bras and Rodriguez-Iturbe (3); and (c) models based on notions of scaling invariance. A brief review of the development of the last type follows.

Lovejoy and Mandelbrot (15) assumed that rainfall scale invariance held in the form of statistical simple scaling (see, e.g., Gupta and Dawdy (6)), specifically, as a *geometric fractal* which will be defined below. Later, Schertzer and Lovejoy (24) were the first to adopt random cascades and multifractal measures, and used them to describe the spatial variability of rainfall. They developed a continuous cascade model, which is somewhat different in details from the discrete cascade which we use in this paper. After Gupta and Waymire (8) explored the MKP function (Mandelbrot (17), Kahane and Peyriere (12)) of several kinds of discrete cascade models, Over and Gupta (21) analyzed the structure of mesoscale rainfall using a random cascade generator by employing the so-called β model. Over and Gupta (22) tested a theory of space-time rainfall applicable to fields advecting without deformation of the coordinates using a random cascade whose generator is a *composite generator* (β model multiplied by the lognormal model).

Jothiyangkoon *et al.* (11) expanded the model of Over and Gupta (22) and introduced a deterministic multiplier estimated by rainfall data into a spatially non-homogeneous composite generator. Tachikawa *et al.* (27) also introduced a deterministic multiplier estimated by elevation, since precipitation depth depends on elevation (Suzuki et al. (25)). On the other hand, Gupta *et al.* (5) and Troutman and Over (28) used random cascades to clarify the interpretation of scaling exponents of event-based flood discharge in cases where multiscaling holds.

In this paper, we analyze the multifractal properties of temporal and spatial precipitation data. A similar type of paper on time series of daily precipitation was published by Svensson *et al.* (26) which examined multifractal properties of temporal rainfall data, though they did not use a cascade model. In particular, we analyze the applicability of the β model to temporal and spatial data, which is an extension of the work of Over and Gupta (21) and Svensson *et al.*(26). The ultimate goal of this research is to find a downscaling method for GCM data and generation of time-series data by average precipitation based on a composite model. However, in this paper we use a simple β model as an analytical tool and explore the temporal and spatial scaling properties in this way.

This paper consists of five sections. In the following section, we present the fundamental theory of a random cascade and fractal and describe the data we use. We analyze temporal data, that is, time series of daily precipitation, in the third section. The fourth section covers our analysis of spatial precipitation data, and consists of two sub-sections. In the first sub-section, we look at whole datasets comprehensively to find the general characteristics as a function of time. We examine each scan of radar data in the second sub-section. The final section presents our conclusions. Although part of this paper is a translation of a published Japanese paper (Kuzuha *et al.*(13)), we also look at the data with the additional perspective contributed by the second author (T. M. Over) who has joined our group.

A BRIEF REVIEW OF THE THEORY OF RANDOM CASCADES AND THE DATA USED

The Theory of Random Cascades

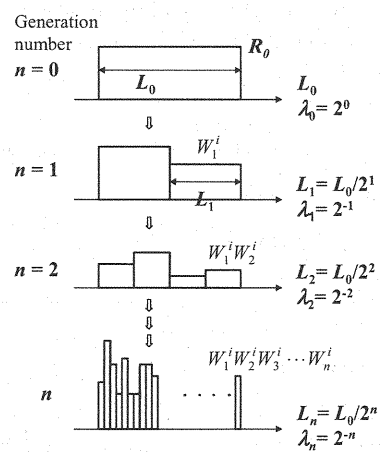


Fig. 1 Schematic of a one-dimensional random cascade.

First, we review the fundamental theory of random cascades, referring to Over and Gupta

(21). Readers who are not familiar with the theory of a (discrete) random cascade can gain a more detailed understanding by referring to Over and Gupta (21, 22) and Over (20). Consider a 'bar' whose length is L_0 on which 'mass' is distributed. Assuming its mass density to be R_0 (mass per unit length), total mass on the 'bar' is R_0L_0 (Fig. 1). A *cascade generator* W divides this total mass into two parts: namely, $R_0L_0 2^{-1}W_1^1$ and $R_0L_0 2^{-1}W_1^2$. The subscript n of W_n^i means the n^{th} generation (see Fig. 1), and the superscript i means the i-th 'bar'. W is a (non-negative) independent identically distributed (IID) random variable, and thus the two masses on the 'bars' are not deterministic. After n iterations of division, the mass on the i-th bar (of n^{th} generation) is $R_0L_0 2^{-n}W_1^iW_2^iW_3^i.....W_n^i$. This example is for one dimension. More general equations for d dimension are as follows. Let us suppose a d-dimensional 'element' instead of a 'bar'. Random cascades are constructed by successively subdividing a d-dimensional element (cube) $[0, L_0]^d$ into b sub-elements, where b is called the branching number. If Δ_n^i denotes the i-th element after n iterations of division (the n^{th} generation), and mass on the element is denoted by $\mu_n(\Delta_n^i)$, then

$$\mu_n(\Delta_n^i) = R_0L_0^d b^{-n} \prod_{j=1}^{n} W_j^i . \qquad (1)$$

The ensemble average of W is defined as follows:

$$\mathrm{E}[W] = 1 \qquad (2)$$

Note that the total mass at each generation is not conserved (however, the average of total mass is equal to initial mass, R_0L_0), since W is an IID random variable (a *canonical* cascade). On the other hand, cascades with pathwise conservation are called *microcanonical* cascades. The *dimensionless spatial-scale parameter* λ_n at the n^{th} generation is defined as follows:

$$\lambda_n = \frac{L_n}{L_0} = b^{-n/d} \qquad (3)$$

where L_n is the side length of the sub-cube (element) at the n^{th} generation. We will apply 1-dimensional random cascades to temporal precipitation data (Section 3) where d = 1 and b = 2, and 2-dimensional ones to spatial data (Section 4). Fig. 1 indicates L_n and λ_n for a 1-dimensional cascade.

The limiting mass $\mu_\infty(\Delta_n^i)$ is obtained by letting n $\rightarrow \infty$, and it satisfies an important recursion equation given by Gupta and Waymire (8),

$$\mu_\infty(\Delta_n^i) = \mu_n(\Delta_n^i)Z_\infty(i) \quad i = 1,2,3.....,b^n \quad , \qquad (4)$$

where $Z_\infty(i)$ are also IID, but we will not discuss this random variable in detail. What is important is 'observed mass' on the element Δ_n^i is $\mu_\infty(\Delta_n^i)$.

For analysis of temporal data under the assumption that it arose from a random cascade, one can easily see that when the branching number is 2, L_0 must be a power of 2. Considering the available data length of daily precipitation at each observation station, we select 8192 (= 2^{13}) (days) ($\cong$ 22.4 (year)) as L_0. $\mu_0(\Delta_0) = R_0L_0$ is the amount of precipitation over 8192 (days), and $\mu_\infty(\Delta_n^i)$ is amount of precipitation over L_n (days). Since we use daily precipitation data, n varies in

the range of 1-13 and L_{13} = 1 (day). In our analysis of spatial data, L_0 is 320 (km). Since the resolution of the spatial data is 2.5 (km) × 2.5 (km) or 5.0 (km) × 5.0 (km) (depending on the observation period mentioned later), the biggest n (generation) is 7 or 6.

It is most straightforward to motivate the notion of scaling invariance in terms of scaling of the moments of the measure. Let us define the qth order moment at level n $M_n(q)$ as follows:

$$M_n(q) = \sum_i [\mu_n(\Delta_n^i)]^q \qquad (-\infty < q < \infty) \tag{5}$$

where $[\mu_n(\Delta_n^i)]^q$ for $q \le 0$ is zero if $\mu_n(\Delta_n^i)$ is non-positive. In view of Eq.(3), the scaling of the moments is defined as

$$\tau(q) = \lim_{\lambda_n \to 0} \frac{\log M_n(q)}{-\log \lambda_n} = \lim_{\lambda_n \to 0} \frac{\log M_n(q)}{(n \log b)/d} \tag{6}$$

Since the masses are identically distributed, if they obeyed the law of large numbers, their sample moments would get close in some sense to their ensemble moments for large n,

$$M_n(q) \approx EM_n(q) = (R_0 L_0^d)^q (b^{1-q} EW^q)^n EZ_\infty^q \,. \tag{7}$$

However, this cannot be made rigorous as it is written, because the ensemble moments are diverging to infinity or going to zero with n depending on the value of $b^{1-q}EW^q$ and taking logs, Eq.(8) is derived as follows.

$$\log_b EM_n(q) = n(1 - q + \log_b EW^q) + \log_b EZ_\infty^q + q \log_b (R_0 L_0^d) \tag{8}$$

Because the scale of resolution λ_n is derived from Eq.(3), that is, $\lambda_n = b^{-n/d}$. Eq. (8) is equivalently written as

$$\log_b EM_n(q) = -d \log_b \lambda_n (1 - q + \log_b EW^q) + \log_b EZ_\infty^q + q \log_b (R_0 L_0^d), \tag{9}$$

which shows that the relationship between the ensemble moments and the scale of resolution exhibits log – log linearity, that is, scale-invariance, with a slope of $d\chi(q)$, where

$$\chi_b(q) \equiv \log_b E[W^q] - (q-1) \tag{10}$$

As for the relationship between $M_n(q)$ and scale, consider the relation

$$\log M_n(q) = \log\left(\frac{M_n(q)}{EM_n(q)}\right) + \log EM_n(q) = \log Y_n(q) + \log EM_n(q). \tag{11}$$

If the ratio $Y_n(q)$ between the moments and the ensemble moments converges for large n, then according to Eq. 11, the moments will scale the same as the ensemble moments, i.e., log-log linearity with scale and a slope of $d\chi(q)$. In fact, a theorem will be cited below to the effect that $Y_n(q)$ does converge to a random variable $Y(q)$, and so the aforementioned scaling of the moments does hold. This theorem further shows that $\tau(q) = d\chi_b(q)$ under certain conditions,

giving the desired result that the moments scale as $\tau(q)$ (Eq. 6).

In this paper, we explicitly consider the β model for which the cascade generator W can be written by using any number p ($0 \leq p \leq 1$):

$$\begin{aligned} W &= 0 \qquad && \text{(with probability } p\text{)} \\ W &= (1-p)^{-1} \qquad && \text{(with probability } 1-p\text{)} \end{aligned} \tag{12}$$

The reader can easily confirm that the expected value of W is unity, as shown in Eq. (2). The means of estimating distribution of W from a data set by estimating the $\tau(q)$ function is based on the following theorem.

THEOREM 2.1 (Holley and Waymire (10)). If $Pr(W<b) = 1$, if there exists a positive number such that $Pr(W > a) = 1$, and if $\dfrac{E[W^{2q}]}{(E[W^q])^2} < b$, then with probability 1

$$\tau(q) = d\chi_b(q) . \tag{13}$$

Furthermore,

$$\lim_{n\to\infty} \frac{M_n(q)}{E[M_n(q)]} = \lim_{n\to\infty} \frac{M_n(q)}{(R_0 L_0^d)^q b^{n\chi_b(q)} E[Z_\infty^q]} \equiv Y(q) \tag{14}$$

where $Y(q)$ is a random variable for each value of q. Equation (14) indicates that $M_n(q)$ does not converge to its expected value in the high-resolution limit due to a random variable $Y(q)$. This demonstrates the non-ergodicity of the measures generated by random cascades (Over and Gupta (21)). Note that $Pr(W > a) = 1$ (a: positive number) does not hold for the β model, because of $W = 0$ having a positive probability. However, a generalized form of this theorem that includes the β model case has since been proven (Molchan (18), Ossiander and Waymire (19)).

In this paper we are mostly concerned with the scaling of the rainy vs. non-rainy areas. This may be computed as follows. Taking $q = 0$ in Eq. (10),

$$\chi_b(0) \equiv \log_b(1-p) + 1 \tag{15}$$

and in accordance with Eq. (13),

$$\tau(0) = d\chi_b(0) = d[\log_b(1-p) + 1] \tag{16}$$

As a result, the following equation holds (Over and Gupta (21));

$$p = 1 - b^{\tau(0)/d - 1} \tag{17}$$

and then p is estimated by using $\tau(0)$, that is, the relationship between $\log M_n(0)$ and $(-\log \lambda_n)$.

The alternative but equivalent method mentioned below is also used below for estimating the p value. Let us define the fractional rainy (wetted) area $f(\lambda_n)$. According to Eqs. (3) and (14), for a large but finite n, the following equation holds (Over and Gupta (21));

$$\begin{aligned} \log M_n(q) &\approx \log Y(q) + q \log(R_0 L_0^d) + \log b^{n\chi_b(q)} + \log E[Z_\infty^q] \\ &= \log Y(q) + q \log(R_0 L_0^d) - \tau(q) \log \lambda_n + \log E[Z_\infty^q] \end{aligned} \tag{18}$$

Since $f(\lambda_n)$ is equivalent to $M_n(0)/b^n$, the fractional rainy area is estimated by

$$\log f(\lambda_n) = \log M_n(0) \approx \log Y(0) + (-\tau(0) + d)\log \lambda_n + \log \Pr[Z_\infty > 0] \quad (19)$$

Estimating the slope (= s) of Eq. (19) ($\log f(\lambda_n)$ versus $\log \lambda_n$), and according to Eq. (16), p is obtained by the following equation (Over and Gupta (21));

$$p = 1 - b^{-s/d} \quad (20)$$

Data

There are two kinds of (surface) meteorological observation systems used by the Japan Meteorological Agency. The first is called AMeDAS (Automated Meteorological Data Acquisition System), which has more than 1,300 observation points; however, observations started relatively recently, after 1976. The other is the 'Surface Meteorological Observation System' which has about 150 observation points and has a relatively long observation period (record length at some observation stations is more than 100 years). Since as mentioned we need a continuous time series for daily precipitation for more than 22.4 years, and there are fewer days when daily precipitation is not recorded in the latter system than the former, we chose the data from the 'Surface Meteorological Observation system' for analysis of temporal precipitation. We attempted to extract the most recent continuous 8192 days of precipitation data at each observation station. However, most time series include 'unavailable days' in which daily precipitation is not recorded. For this reason, we relax the condition for extraction: namely, the latest 8192 days of precipitation in which there are no continuous 64-day periods in which there are more than 2 unavailable days. If there is one unavailable day in the continuous 64 (days), we assume that there is no precipitation on that day. Under these conditions, we were able to extract continuous 8192 daily precipitation data (time series) at 151 observation stations.

For spatial precipitation data, we adopt the 'Radar-AMeDAS Precipitation' data collected by the Japan Meteorological Agency, for which the radar estimates of precipitation rates are corrected by the AMeDAS data. The correction has been carried out due to errors of radar data arising from the instability of the sensitivity of the radar hardware and to the vertical variation of rainfall (Makihara (16)). The rain depths in this data set are not continuous and the resolution is 1.0mm in the range of 1-77mm, 5.0mm in the range of 80-1250mm, and so on. However, in this paper, we regard the discrete value as real rain rate. We use the data for 2001, during which the spatial resolution of the data is 5.0 km × 5.0 km between January and March, and 2.5 km × 2.5 km between April and December.

SCALING OF TEMPORAL DATA

Scale Invariance of Temporal Data

Fig. 2 shows examples of the relationships between $\log_2 \lambda_n$ and $\log_2 M_n(q)$ at Wakkanai (Local ID number 401; WMO ID number 47401) and Izuhara (800, 47800). As mentioned above, the relationship between $\log_2 \lambda_n$ and $\log_2 M_n(q)$ should show linearity if the time series of daily precipitation has scaling or fractal properties. Svensson *et al.* (26) found that daily precipitation data from China and Sweden have scaling properties in the range of 1-32 days. We carried out the same preliminary investigation for $0 \le n \le 13$ (corresponding to 8192 days $\ge L_n \ge$ 1 day) and

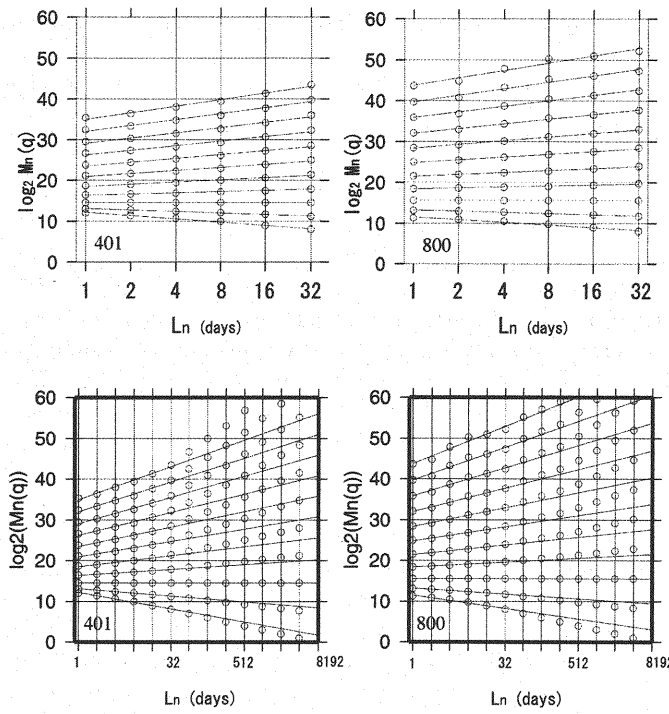


Fig. 2 Top row: The relationship between L_n and $\log_2 M_n(q)$ at Wakkanai (Local ID number 401; WMO ID number 47401) and Izuhara (800, 47800). Circles are data and lines are regression lines determined by least squares, and the data and lines correspond to q = 0, 0.5, 1.0, 1.5,....5.0 from the bottom. Bottom row: The same figures as the top two figures, but the range of L_n is wider. One can confirm that the relationship between L_n and $\log_2 M_n(q)$ is almost linear for shorter L_n; say, less than 32 days.

obtained the same conclusion. Thus we show the relationships in the range of 1 to 32 days in Fig. 2; other figures which are not shown here show mostly same results. The analysis made by Svensson *et al.* (26) was of monsoon climate data in China and temperate climate data in Sweden. Both regions have a temperate climate, according to Köppen's regional classification, but the former is affected by monsoons, Baiu cold fronts and typhoons. They showed two examples of the relationship between $\log_2 \lambda_n$ and $\log_2 M_n(q)$ for a monsoon climate (similar to Fig. 2 in this paper). One is at a location which was affected by a major typhoon in 1975, and the other is for a location which was not so affected. Curves for the former exhibit well-defined straight line behavior for $q \leq 2.5$, but the moments of high order do not follow a straight line due to an exceptionally high 5-day rainfall caused by the typhoon, while curves for the latter and for the Swedish locations exhibit linearity for $0 \leq q \leq 4.0$. Although Japan suffers from numerous typhoons, there has been no extraordinary precipitation, such as that seen in China in 1975, in the dataset used in this paper, and thus the relationship exhibits a relatively well-defined linearity. However, the relationship generally departs slightly from linearity for large values of q.

Rigorous statistical work on the significance of deviations from linearity in the relationship

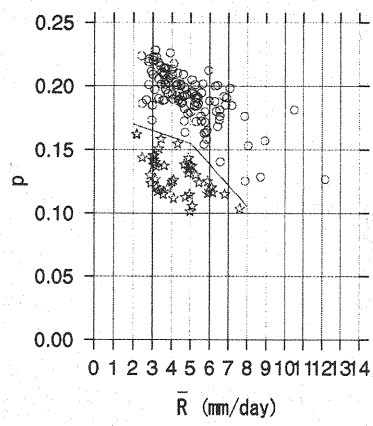


Fig. 3 The relationship between $\bar{R}$ and p for 151 daily precipitation observation stations. The line in the figure divides the set of data into the two groups. Circles (above the line) denote the data along the Pacific Ocean (Group A) and stars (below the line) denote those of the data along the Japan Sea (Group B).

between $\log M_n(q)$ and $\log \lambda_n$ has been carried out by Troutman and Vecchia (29) and Ossiander and Waymire (19) for theoretical cascades in which the measure is simulated according to the boxes shown in Fig.1 (an 'on-grid' cascade in the terminology of Over (20)). However, of course real data does not follow these boxes (thus, it is an 'off-grid' cascade; see APPENDIX), and so these results can not be applied directly to real data. Adaptation of these results to real data is beyond the scope of this paper. As deviations from log-log linearity are obvious to the eye for time scales longer than 32 days, we conclude, preliminarily, on the basis, as others have done, that scale-invariance is satisfied in the range 1 to 32 days but not at longer time scales.

Estimation of p Value of β Model

For each series of 8192 daily values, we calculated the p value using $\tau(0)$ according to Eq. (17), where $\tau(0)$ was computed according to the method shown in Fig. 2 over the range of finite scales from 1 to 32 days. We further calculated the average daily precipitation ($\bar{R}$ in mm/day) for each time series. Fig. 3 indicates the relationship between $\bar{R}$ and p for 151 stations. As can be seen from Eq.(12), although p does not directly represent the fraction of dry periods, p is the ratio by which dry periods propagate. Thus it follows that if day periods at the time scale where the cascade apparently begins (32 days) are the same, then the larger p is, the more dry days there will be in the record. If we further assume that the daily rainfall depth on days when there is rain is independent of p, then the larger p is, the smaller $\bar{R}$ should be. However, the data in Fig. 3 does not indicate such a relationship. We divide the set of points into two groups using a broken line in the figure; namely, above the line and below the line. By dividing in this manner, a decreasing tendency of p with $\bar{R}$ becomes clear in each group. Hereinafter, the group above the broken line

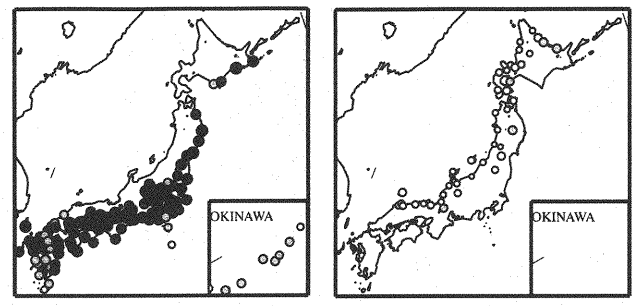


Fig. 4 Distribution of surface meteorological observation stations: (a) Group A, (b) Group B. The diameter and density of color of each circle indicates the value of p. The diameter of each circle is proportional to the value of p, and the darkness of color in the circle indicates the class of value of p; there are four classes of thickness: $p \leq 0.15$, $0.15 < p \leq 0.18$, $0.18 < p \leq 0.2$, and $0.2 < p$, respectively.

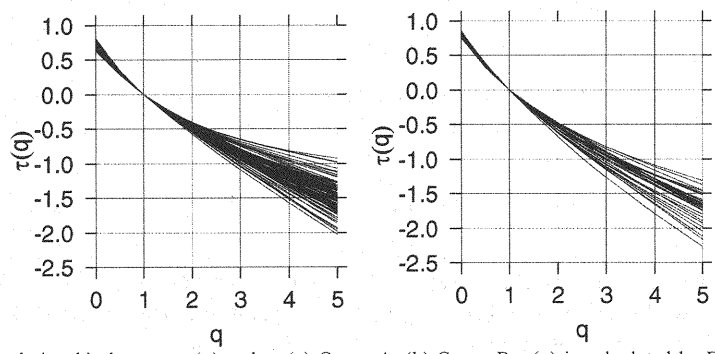


Fig. 5 The relationship between $\tau(q)$ and q: (a) Group A, (b) Group B. $\tau(q)$ is calculated by Eq. (6) using least squares regression over the range of scales $1 \leq L_n \leq 32$.

is referred as 'Group A' and below the line as 'Group B'. Fig. 4 shows the distribution of 151 the observation stations; the left-hand figure is for Group A and the right-hand figure is for Group B. One can find that 'Group A', with a relatively large p, corresponds to observation stations along the Pacific Ocean, and 'Group B', with a relatively small p, corresponds to observation stations along the Japan Sea. This means that the scaling properties of daily precipitation along the Pacific Sea are different from those along the Japan Sea. The former is more intermittent, and the latter less so. We hypothesize that intermittent storms in summer and fall (which are dominant along the Pacific) and continuous snowfall in winter (which mainly falls in the region along the Japan Sea) cause these properties.

Multifractal Properties

According to Frisch and Parisi (4), the relationship between $\tau(q)$ and q exhibits a convex curve if data has multifractal properties, while the relationship exhibits a linear pattern if data is monofractal. According to Over and Gupta (21), however, the relationship exhibits linearity in the case of the β model, which is a multifractal model; that is, the relationship exhibits linearity if data can be completely modeled using only the β model. Thus, we designate as *multifractal in the narrow sense* cases that are strictly convex, and *multifractal in the wide sense* cases including β models which are multifractal models but whose $\tau(q)$ versus q exhibits linearity. Fig. 5 gives the $\tau(q)$ curves for the two groups of daily data. This figure includes only data at stations at which r^2

of $\tau(q)$ versus q for $q = 5$ is greater than 0.95 and were computed by fitting over the range of scales of 1 to 32 days. Fig. 5 shows that 'Group A' generally exhibits more convexity than 'Group B', indicating that 'Group A', or stations along the Pacific, have stronger multifractal properties and intermittent and peaky properties than stations along the Japan Sea. As mentioned above, a time series which can be completely modeled using the β model must exhibit linearity in $\tau(q)$ vs. q. (compare Fig. 5). Thus, we can conclude that using only the β model is inadequate for generating / modeling precipitation time series in this region and over these scales. However, the evaluation of p values in this section is meaningful, as it depends only on the structure of the dry and rainy periods (i.e. whether wet or dry), not the values during wet days (i.e. precipitation amount during wet period) , which cause the convexity in the $\tau(q)$ curves.

SCALING OF SPATIAL DATA

Questions to be Addressed

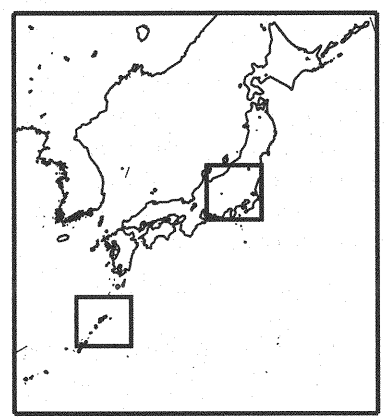


Fig. 6 The locations of a land region (box in lower left) and a sea region (box in upper right).

Over and Gupta (21) used a β model to analyze spatial precipitation data and investigated the dependence of the parameters of the β model on large (meso) scale forcing measured by the spatial average of the precipitation field. They used data from GATE (Atlantic Tropical Experiment) Phase I and II. These data were obtained from C-band radar images over the tropical Atlantic during the summer of 1974. Phase I was between 28 June and 16 July, and Phase II was between 28 July and 15 August. Our present investigation was motivated by the following three questions.
1) Although Over and Gupta (21) analyzed using the complete data set as well as each scan, data length is 40 days at most. What are the scaling properties for more long-term data?
2) What is the difference between data for the summer and winter?
3) Can the random cascade theory be applied to spatial data on land, where topography, especially the existence of mountains, may disturb multifractal or scaling properties?

We extracted data for a land region and a sea region between January and February and between July and September in 2001 from the complete radar dataset for the purpose mentioned above. The land region is located between 34.5°N to 37.7°N and 137°E to 141°E, and the sea region is located between 26.5°N to 29.7°N and 127.5°E to 131.5°E (Fig. 6). Both regions are 320 × 320 km; i.e., L_0 is 320 km. Recall that for January and February the highest resolution is 5.0 km,

yielding a 64 by 64 pixel grid over the 320km, while for July through September, the resolution is 2.5 km, yielding a 128 by 128 pixel grid. In this section, we analyze the scaling properties using the whole dataset (which consists of data for January, February, July, August and September), and then analyze it using each scan of spatial precipitation. Radar AMeDAS precipitation data is an accumulation of hour of precipitation; thus, each of our datasets consists of 3624 (151 day × 24 hour) scans of data. However, since radar is not always available and there are some scans which include too few rainy pixels, preventing regression analysis from being carried out, 2302 scans are available for the land region and 2717 scans are available for the sea region.

Analyses Using Whole Dataset

First, we calculate the p value for each scan of data of 5019 (land: 2302, sea: 2717) scans. The p value is calculated using Eq. (20) by way of the slope of $\log\lambda_n$ versus $\log f(\lambda_n)$. $f(\lambda_n)$ is the fractional rainy area. To understand this function, let us assume that there is only one pixel where there is precipitation. Even if average precipitation is very small in this particular case, $f(\lambda_0)$ is 1 and $\log f(\lambda_0)$ is 0. Thus one can find that $\log f(\lambda_n)$ can easily be zero for small values of n (i.e., at large scales). Fig. 7 shows two examples of the relationship. In this figure natural logarithms are used. The regression analysis in the figure was carried out by least squares weighted by the number of pixels, after Gupta and Waymire (8). It can be observed that log-log linearity holds for large but finite values of n. Note that the intercept of the regression line is negative in the left-hand figure, while it is positive in the right-hand figure. Over and Gupta (21) have clarified the meaning of the intercept. According to Eq. (19), the intercept corresponds to $\log Y(0) + \log \Pr[Z_\infty > 0]$; and they mentioned that $\log \Pr[Z_\infty > 0]$ is negative but small in magnitude for p = 0.21~0.35; thus the intercept corresponds approximately to $\log Y(0)$. According to Eq. (14),

$$\lim_{n\to\infty} \frac{M_n(0)}{E[M_n(0)]} = \lim_{n\to\infty} \frac{f(\lambda_n)}{E[f(\lambda_n)]} = Y(0) \qquad (21)$$

This means that a positive, large intercept corresponds to an unusually (where 'usually' is defined in terms of scale invariance holding over all scales) large fractional rainy area (or, equivalently concavity in the $\log f(\lambda_n)$ versus $\log\lambda_n$ curve) and that a negative, small intercept corresponds to an unusually small fractional rainy area (or, equivalently convexity in the $\log f(\lambda_n)$ versus $\log\lambda_n$ curve). Thus the left-hand figure in Fig. 7 shows a typical case of an unusually small fractional rainy area and the right-hand figure shows an unusually large fractional rainy area. One can see that in the right-hand figure in Fig. 7, $\log f(\lambda_n)$ is zero for n = 0,1(320 km and 160 km). This is typical for the case of an unusually large fractional rainy area. Since rain areas are spread widely, after the 1st division by random cascade, all four of the elements possess precipitation ($f(\lambda_n) = 1$ and $\log f(\lambda_n) = 0$).

Fig. 8 indicates the time variation of the intercept and the p value for the data from the two

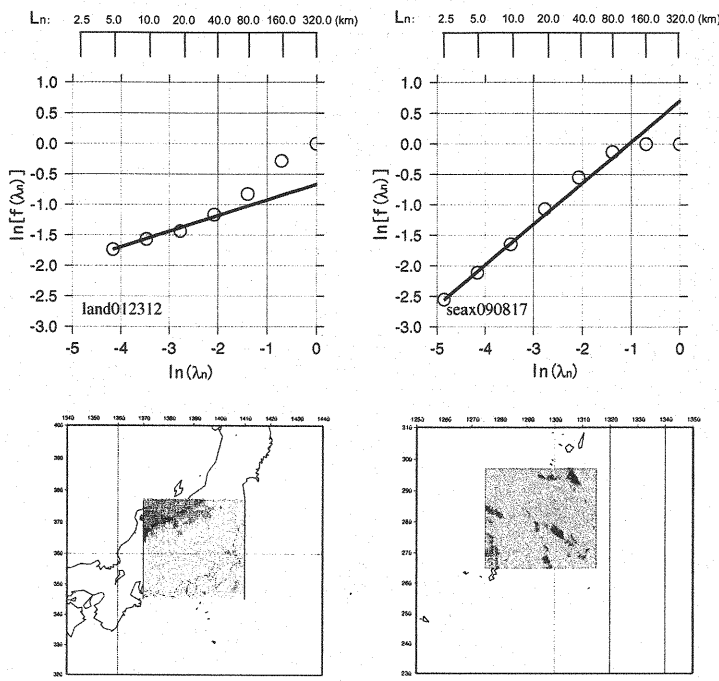


Fig. 7 Top row: Examples of the relationship between $\log\lambda_n$ (= - n) and $\log f(\lambda_n)$, determined by weighted least squares regression: (a) the land region on January 23, (b) the sea region on September 23. Bottom row: Spatial distribution of rainfall intensity in the scenes analyzed.

regions. For the sea region, the positive intercept is dominant, especially in the summer, while in the land region, the negative intercept is slightly dominant especially in the latter part of the year. Examining the sea region in the summer further, notice the p value is also large. This means that the structure of rainy and non-rainy area is like that shown in Fig.7(b), and thus the $\log f(\lambda_n)$ curve is concave, slope of $\log f(\lambda_n)$ vs. $\log\lambda_n$ is large in small scales, and consequently p is large. This implies that non-rainy areas at large scales are not as many as the small-scale scaling structure would imply. On the other hand, positive intercept may also arise due to the effect of the data being 'off-grid'(Over (20); see APPENDIX in this paper), and in the present dataset, from the effect of advection over the hour of accumulation of the radar data. Over and Gupta (21) mentioned that this unusually large fractional rainy area is typical in the GATE data. Our conclusion is the similar to theirs, since they investigated only a sea region. Moreover, we found that this is not typical on land. Fig. 7 also suggests that scale invariance does not hold for small values of n (large spatial scales). This finding was also observed by Gupta and Waymire (8) and Over and Gupta (21), with a critical scale of about 100 km.

Fig. 9 indicates the relationship between estimates of p values and spatial average of precipitation ($\bar{R}$) for the whole dataset. As might be expected, there is not as strong a one-to-one relationship between the parameter of the β model and large-scale forcing measured by the spatial average precipitation as was found by Over and Gupta (21) and Over and Gupta (22), because the relationship in the figure includes precipitation caused by various kinds of precipitation mechanisms. Thus we will further investigate the existence of a one-to-one functional relationship

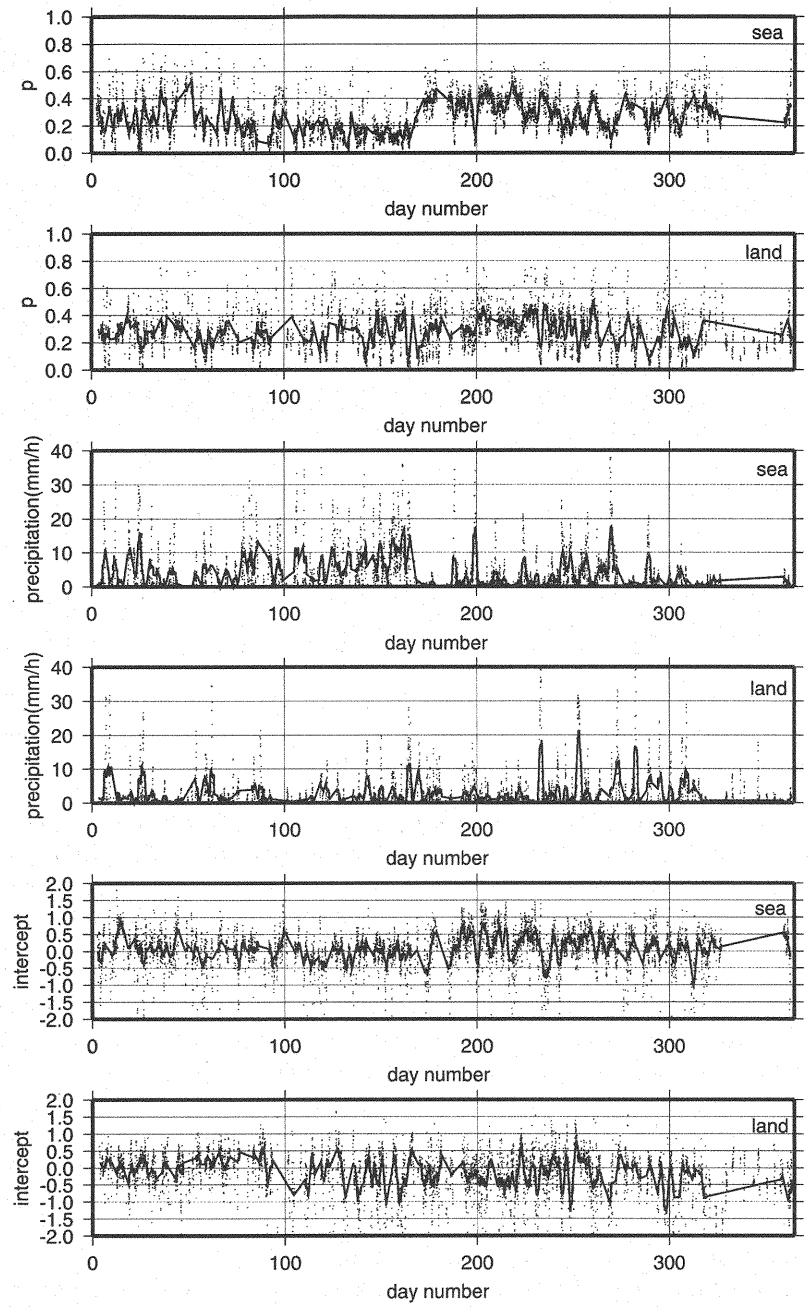


Fig. 8 Time variation of intercept, average of precipitation and p value for the land region and a sea region: (a) the p value in the sea region, (b) the same in the land region, (c) the average of precipitation in the sea region, (d) the same in the land region (e) the intercept in the sea region, (f) the same in the land region. In these figures, annual series data are shown. Dots indicate hourly data and solid lines indicate a 49 -hour moving average.

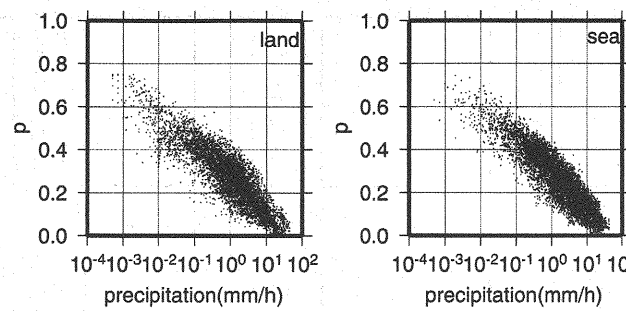


Fig. 9 The dependence of p value on spatial average precipitation for whole dataset: (a) the land region, (b) the sea region.

by extracting separate precipitation events later.

Analyses Using Each Scan of Radar Data

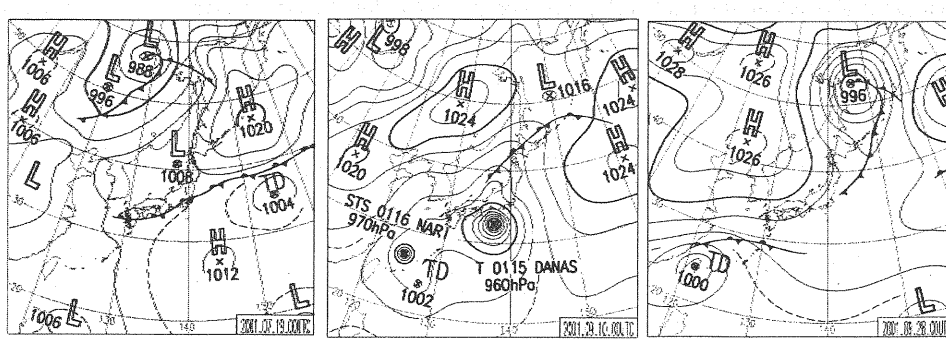
Here we use several scans of radar data and investigate one effect of mesoscale forcing (spatial average of precipitation) on the single parameter of the β model like Over and Gupta's (21). We need to classify mechanisms that are causes of precipitation and investigate the effects of making these classifications. Since we have more than 5,000 radar data scans, however, we regard this as an open problem and have selected the following four independent precipitation events (refer to weather charts made by JMA in Fig. 10):

(1) July 18 to 20 (sea region), (2) September 9 to 11 (land region), (3) September 27 to 29 (sea region) and (4) September 27 to 29 (land region).

(1) is a period when there is no front over the sea region, while there is a stationary front over the Kyushu district whose location is north of the sea region. According to Radar AMeDAS data, a strong precipitation area is moving north-eastward over the region.

(2) is a period when a typhoon passes over the land region.

(3) and (4) show a period when there is a stationary front in the northern part of the sea region and there is a typhoon in the west of the sea region. According to the weather chart, there is no front over the land region.



(a) (b) (c)

Fig. 10 Weather charts around Japan. (a) case (1) at July 19 00UTC, (b) case (2) at September 10 00UTC, and (c) cases (3) and (4) at September 28 00UTC. Note that JST is 9 hours ahead of UTC.

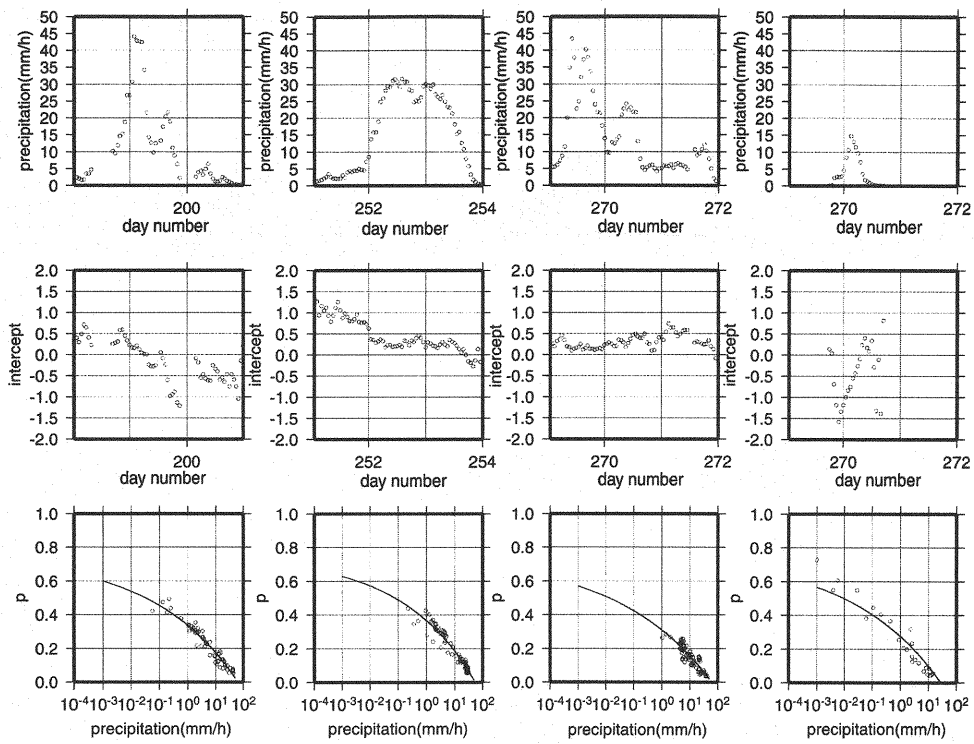


Fig. 11 Time variation of spatial average of precipitation (top row) and intercept of regression line (second row), and the relationship p and $\bar{R}$ (bottom row). Columns 1, 2, 3 and 4 correspond to cases (1), (2), (3) and (4), respectively. The day number is counted from January 1. In the figures in the botton row, regression lines are indicated. p is estimated by Eq. (20).

Fig. 11 indicates the time variation of spatial average of precipitation ($\bar{R}$), the intercept mentioned above, and the relationship between p and $\bar{R}$ for these four events. Over and Gupta (21, 22) reported two different functions of p and $\bar{R}$ as follows:

$$\left(\frac{\bar{R}}{R_{\max}}\right)^{s} = 1 - \frac{p}{0.75} \tag{22}$$

$$p = \log_{10}\left(\frac{\bar{R}}{R_{\max}}\right)^{s} \tag{23}$$

where s and R_{max} are parameters. We applied both the equations to the data and the result of the regression analyses are shown in Table 1. Equation (22) (Over and Gupta (21)) seems to be slightly more appropriate for our data. The solid line in the figure is the regression curve provided by the equation. However, of course, we cannot conclude which equation is generally appropriate.

Table 1 Estimates of the parameters of Eq. (22) and Eq. (23) and r^2.

	Equation (22) (Over and Gupta (94))			Equation (23) (Over and Gupta (96))		
	s	R_{max}	r^2	s	R_{max}	r^2
Period (1)	0.145	62.1	0.9235	-0.166	102	0.9118
Period (2)	0.167	55.6	0.9277	-0.198	77.2	0.9117
Period (3)	0.129	60.8	0.7737	-0.180	71.1	0.7746
Period (4)	0.136	28.0	0.9452	-0.162	32.2	0.9444

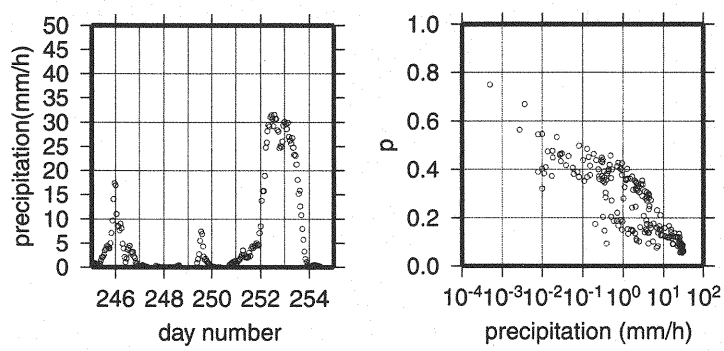


Fig. 12 Time variation of spatial average of precipitation (left) and the relationship between p and $\bar{R}$ (right). This figure allows inference of a close relation between single-peaks of precipitation and a one-to-one relationship between p and $\bar{R}$. This figure is an example that there are multi-peaks of precipitation. (2001/09/03/0:00 - 2001/09/12/24:00)

We carried out an examination based on data from other precipitation events, and conclude that the relationship between $\bar{R}$ and p is defined by a one-to-one function if each independent precipitation event is extracted, and even if there is a front over the region. Extracting each independent precipitation event from a major convective storm is relatively easy (e.g. a typhoon), but extracting individual events from data which includes multiple precipitation peaks, and where the time variation of precipitation is complex, is more difficult. If we extract time varying precipitation which includes several precipitation peaks, it seems that there is a possibility that we cannot assume a one-to-one function which represents the relationship (see. Fig.12, which is an example whose relationship between $\bar{R}$ and p does not exhibit one to one function, and there are multi-peak of precipitation).

CONCLUSIONS

We examined temporal and spatial precipitation data and used β model to analyze both sets of data. Daily temporal data has scaling properties in the range of 1 day ≤ L_n (the duration of

accumulation of daily precipitation) ≤ 32 (days). Although scaling invariance does not hold for time series longer than this, the p value, the parameter of the β model, can be used as an index to represent the scaling properties of the dataset. Based on this parameter, we conclude that precipitation at surface observation stations along the Pacific Ocean is more intermittent and peaky than those along the Japan Sea. Moreover, although both groups have multifractal properties, those along the Pacific exhibit more multifractality.

As for the spatial data, we analyzed the time variation of scaling properties using the intercept of regression analysis. These results showed that over the sea in the summer, there was a tendency for a different scaling as small vs. large scales, which may be due to several causes: the 'off-grid' nature of the scale, the accumulation of rain depth over one hour, or actual deviations from scaling invariance.

Finally, we extracted four radar data scans and examined whether or not a one-to-one function can represent the relationship between large-scale forcing (spatial average of precipitation) and the parameter of the β model. We conclude that there is a one-to-one function provided we extract individual precipitation events, that is, those which do not contain multiple peaks. The equation relating p and the spatial average rain rate in Over and Gupta (21) is slightly more appropriate than in Over and Gupta (22).

ACKNOWLEDGEMENT

The authors are grateful to Vijay K. Gupta (University of Colorado) for his valuable advice on scaling, fractals and the theory of random cascades.

REFERENCES

1. Austin, P. and R.A. Houze: Analysis of stracture of precipitation patterns in New England, Journal of Applied Meteorology, 11, pp.926-935, 1972.
2. Blöschl, G. and M. Sivapalan: Process controls on regional flood frequency: Coefficient of variation and basin scale, Water Resources Research, 33, pp.2967-2980, 1997.
3. Bras, R.L. and I. Rodriguez-Iturbe, 1976: Rainfall generation: a nonstationary time-varying multidimensional model. Water Resources Research, 12(3), 450-456.
4. Frisch, U. and G. Prisi: On the singularity structure of fully developed turbulence, in Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics, edited by M. Ghil, R. Benzi, and G. Parisi, pp.84-88, 1985.
5. Gupta, V.K., S.L. Castro and T.M. Over: On scaling exponents of spatial peak flows from rainfall and river network geometry, Journal of Hydrology, 187, pp.81-104, 1996.
6. Gupta, V.K. and Dawdy: Regional analysis of flood peaks: multiscaling theory and its physical basis. In: Advances in Distributed Hydrology, ed. R. Rosso, A.Peano, I.Becchi, and G.A.Bemporad, pp.149-168, 1994.
7. Gupta, V.K. and Dawdy: Physical interpretations of regional variations in the scaling exponents of flood quantiles, Hydrological Process, 9, pp.347-361, 1995.
8. Gupta, V.K. and E.C. Waymire: A statistical analysis of mesoscale rainfall as a random cascade, Journal of Applied Meteorology, 12, pp.251-167, 1993.
9. Gupta, V.K. and E.C. Waymire: Spatial Variability and scale invariance in hydrologic regionalization, In: Scale Dependence and Scale Invariance in Hydrology, ed. G. Sposito, pp. 88- 135, 1998.
10. Holley and Waymire: Multifractal dimensions and scaling exponents for strongly bounded random cascades, Annals of Applied Probability, 2, pp.819-845, 1992.

11. Jothiyangkoon C., M. Sivapalan and N.R. Viney: tests of a space-time model of dairy rainfall in southwestern Australia based on nonhomogeneous random cascades, Water Resources Research, 36, pp.267-284, 2000.
12. Kahane,J.P. and J. Peyriere: Sur certaines martingales de Benoit Mandelbrot, Advances in Mathematics, 22, pp.131-145, 1976.
13. Kuzuha Y. K. Tomosugi and T. Kishii, Analyses of temporal and spatial precipitation data by random cascades, Annual Journal of Hydraulic Engineering, 47, pp.133-138, 2003. (in Japanese)
14. LeCam, L.: A stochastic description of precipitation, Proceedings, Fourth Berkeley Symposium on Mathematical Statistics and Probability, 3, Berkeley: University of California, pp.165-186, 1961.
15. Lovejoy, S. and B.B. Mandelbrot: Fractal properties of rain and a fractal model, Tellus, 37A, pp.209-232, 1985.
16. Makihara,Y.: A Method for Improving Radar Estimates of Precipitation by Comparing Data from Radar and Rain gauges, Journal of the Meteorological Society of Japan, 74, pp.459-480, 1996.
17. Mandelbrot, B.B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, Journal of Fluid Mechanics, 62, pp.331-358, 1974.
18. Molchan, G.M.: Scaling exponents and multifractal dimensions for independent random cascades, Communications in Mathematical Physics, 179, pp.681-702, 1996.
19. Ossiander, M. and E.C. Waymire, Statistical estimation for multiplicative cascades, Annals of Statistics, 28, pp.1533-1560, 2000.
20. Over, T.M.: Modeling space-time rainfall at the mesoscale using random cascades, Ph.D. thesis, University of Colorado, 1995.
21. Over, T.M. and V.K. Gupta: Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing, Journal of applied meteorology, 33, pp.1526- 1542, 1994.
22. Over, T.M. and V.K. Gupta: Space-time theory of mesoscale rainfall, Journal of Geophysical Research, 101, pp.26319-26331, 1996.
23. Robinson and M. Sivapalan: An investigation into the physical cause of scaling and heterogeneity in regional flood frequency, Water Resources Research, 33, pp.1045-1059, 1997.
24. Schertzer, D. and S. Lovejoy: Physical modelingand analysis of rain and clouds by anisotropic scalingmultiplicative processes, Journal of Geophysics Research, 92, pp.9693-9714, 1987.
25. Suzuki, Y., E. Nakakita and S. Ikebuchi: Study on the dependence properties of rainfall distributions on topography based on the altitude dependence line, Annual Journal of Hydraulic Engineering, 47, pp.301-307, 2001. (in Japanese with English abstract)
26. Svensson, C., J. Olsson and R. Berndtsson, Multifractal properties of daily rainfall in two different climates, Water Resources Research, 32, pp.2463-2472, 1996.
27. Tachikawa, Y, M. Hiwasa and K. Takara: Spatial rainfall field simulation by using random cascade and dependence line on topographic elevations, Annual Journal of Hydraulic Engineering, 47, pp.127-132, 2003. (in Japanese with English abstract)
28. Troutman, B.M. and T.M. Over: River flow mass exponents with fractal channel networks and rainfall, Advances in Water Resources, 24, pp.967-989, 2001.
29. Troutman, B.M. and A.V. Vecchia: Estimation of Renyi exponents in random cascades, Bernoulli, 5, pp.191-207, 1999.

APPENDIX

'Off-grid' here refers to the fact that the theory and simulation of discrete random cascades are done with reference to a regular grid of boxes that fill a cube in d-dimensional space, whereas, in general, data will not honor this grid. Therefore, for example, a compact region of rainfall of about the same area as a box in this grid at some scale will likely cover portions of two or more boxes, and the fraction of boxes at this scale that have non-zero rain will be, in some sense, artificially enlarged. Over (20) argues further that this effect is not scale-invariant and introduces concavity into the scaling of the $q = 0$ moment, though not enough to explain the observed deviations from scale-invariance observed in the GATE data examined there. An 'off-grid' effect would also exist for $q > 0$ but would be smaller.

(Received August 25, 2003 ; revised November 26, 2003)