Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond
by Jianbo Gao, Yinhe Cao, Wen-wen Tung and Jing Hu

CHAPTER 10

STAGE-DEPENDENT MULTIPLICATIVE PROCESSES

We have discussed three different types of fractal models: $1/f^{\beta}$ processes with long memory, stable laws and Levy motions, and random cascade multifractals. To see better the relations among these models, in this chapter we describe a stage-dependent multiplicative process model.

10.1 DESCRIPTION OF THE MODEL

The stage-dependent multiplicative process is best described in a recursive manner (see the schematic of Fig. 10.1). It is a conservative model and can be readily modified to be a nonconservative model.

Step 1: At stage 0, we have a unit interval and a unit mass (or weight). The unit time interval is the total time span of interest. It could be 1 min, 1 hr, or even 1 year. The unit mass could be the total energy of a turbulent field under investigation or the total traffic loading to a network in the entire time span. We divide the unit interval into two (say, left and right) segments of equal length, partition the mass into two fractions, $r_{1,1}$ and $1 - r_{1,1}$, and assign them to the left and right segments, respectively. The first subscript "*1*" is used to indicate the stage number 1. The second subscript "1" denotes the

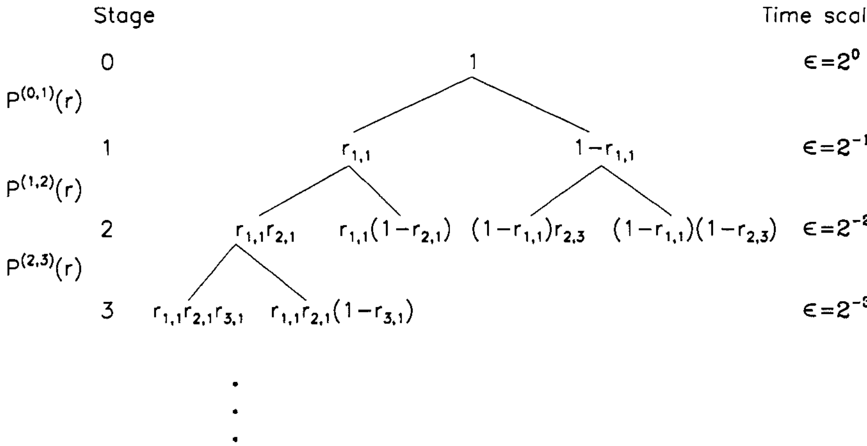

Figure 10.1. Schematic illustrating the construction rule of the stage-dependent multiplicative process model.

position of the weight at that stage, starting from the left. We only assign odd positive integers, leaving the even integers to denote positions occupied by the "complementary" weights such as $1 - r_{1,1}$. The parameter $r_{1,1}$, called the multiplier, is a random variable governed by a PDF $P^{(0,1)}(r)$, $0 \le r \le 1$, where the superscript "(0,1)" is used to indicate the transition from stage 0 to stage 1. $P^{(0,1)}(r)$ is assumed to be symmetric about $r = 1/2$, so that $1 - r_{1,1}$ also has the distribution $P^{(0,1)}(r)$.

Step 2: Divide each interval at stage i into two segments of equal length. Also, partition each weight, say $r_{i,1}$, at stage i, into two fractions, $r_{i,1}r_{i+1,1}$ and $r_{i,1}(1 - r_{i+1,1})$, where $r_{i+1,1}$ is a random variable governed by a PDF $P^{(i,i+1)}(r)$, $0 \le r \le 1$, which is also symmetric about $r = 1/2$. Note that the random variables, $r_{i,j}$'s, are independent of all the other random variables used in the multiplicative process.

Step 3: A complex time series may be modeled with arbitrary precision by simply making the variances stage independent for a suitably chosen parameterized density function $P^{(i,i+1)}(r)$. However, while modeling a time series of length n, this would amount to estimating $\log n$ parameters, and given the universality result, one would suffer from all the problems inherent in global optimization problems, such as overfitting and being stuck in local minima. Hence, a more relevant question is how to constrain the parameter space and still retain the ability to model different types of stochastic processes. We require that the ratio of the variances satisfy a simple constraint,

$$\sigma^2(i, i + 1) = \gamma\sigma^2(i - 1, i), \tag{10.1}$$

for a constant $\gamma \geq 1$. As we show in our analysis, it is indeed the case that by varying the single parameter γ, one can generate a wide range of processes including processes with long memory, pure multiplicative multifractals, and SRD processes. The implications of these results are further elaborated upon in Sec. 10.3.

Step 4: Interpret the weights at stage N, $\{w_j, j = 1, \cdots, 2^N\}$, as the process being modeled (such as a counting traffic process).

By comparing the above model with the pure multiplicative multifractal model described in Chapter 9, we immediately obtain the following properties of the model:

- When the PDFs $P^{(i,i+1)}(r)$ are stage independent, the model reduces to the ideal multiplicative multifractal model. This is the case when $\gamma = 1$ and all the $P^{(i,i+1)}(r)$ are of the same functional form.

- When $P^{(i,i+1)}(r)$ are of the same functional form but γ is slightly larger than 1, the process is not multifractal. However, the weights at stage N will have an approximately log-normal distribution.

- The variance parameter $\sigma^2(i, i+1)$ describes the burstiness of the process at the particular time scale $2^{-(i+1)}T$, where T is the total time span of interest. Hence, when $\gamma = 1$, the burstiness level remains unchanged for all time scales. When $\gamma > 1$, the burstiness level of the process decreases from shorter to longer time scales.

In the next section, we shall show that the γ parameter and the Hurst parameter are related by a simple relation (when $\gamma > 1$; the case of $\gamma = 1$ has been dealt with in Chapter 9),

$$2^{2-2H} \approx \gamma. \tag{10.2}$$

Hence, when $\gamma = 2$, the decrease in the burstiness level of the process is such that the process effectively becomes a short-range-dependent process such as a Poisson process.

We now comment on Eq. (10.1). Suppose that $\sigma^2(0, 1) > 0$ is given. Then $\sigma^2(i, i+1) \to \infty$ when $i \to \infty$. This contradicts the restrictions made on the random variables, i.e., any distribution is nonzero only between 0 and 1 and is symmetric about $1/2$, and $0 \leq \sigma^2(i-1, i) \leq \frac{1}{2}$ for any i. The contradiction can be resolved by noticing that after a certain stage i^*, the stage-dependent multiplier distributions can no longer change with i (for example, if the shape of the multiplier distribution is initially Gaussian, when the variance gets larger and larger, within the unit interval the distribution eventually settles to the uniform distribution). Beyond stage i^*, the model reduces to the cascade multifractal discussed in Chapter 9. In this chapter, we shall focus on the time scales above that of stage i^*.

10.2 CASCADE REPRESENTATION OF $1/F^{\beta}$ PROCESSES

Since $E(w_j) = 2^{-N}$, the process, $\{w_j - 2^{-N}, j = 1, \cdots, 2^N\}$, is thus a zero mean time series. Below we shall prove that the random walk process, defined by

$$y_n = \sum_{j=1}^{n} (w_j - 2^{-N}), \tag{10.3}$$

is a fractal process.

Theorem 1 The random walk process y_n is asymptotically self-similar, characterized by a power spectral density $E(f) \sim f^{-(2H+1)}$, with $2^{2-2H} \approx \gamma$, when $\gamma > 1$.

Proof First, we note that a $1/f^{(2H+1)}$ process is characterized by

$$F(m) = E[(y_{n+m} - y_n)^2] \sim m^{2H}. \tag{10.4}$$

Hence,

$$F(2m)/F(m) = 2^{2H}. \tag{10.5}$$

It thus would suffice for us to prove that the ratio between the variance for the weights at stage $i-1$ and stage i is given by $4/\gamma$. A weight at stage i can be written as $r_1 r_2 \cdots r_i$, where r_k, $k = 1, ..., i$, is governed by a PDF $P^{(k-1,k)}(r)$. Thus, the variance for a weight at stage i is

$$var(i) = [1/4 + \sigma^2(0,1)][1/4 + \gamma\sigma^2(0,1)] \cdots [1/4 + \gamma^{i-1}\sigma^2(0,1)] - 2^{-2i}.$$

For simplicity, let us define

$$c^2 = \frac{\sigma^2(0,1)}{(1/2)^2}.$$

Hence,

$$var(i) = (\frac{1}{2})^{2i}\{\Pi_{j=1}^{i}(1 + \gamma^{j-1}c^2) - 1\}.$$

Since any $P^{(k-1,k)}(r)$ is defined in the unit interval and is symmetric about $r = 1/2$, its variance must be in the interval $[0, 1/4]$. This means that

$$\gamma^{k-1}c^2 < 1$$

for any k. Hence,

$$var(i) \approx (\frac{1}{2})^{2i}(\gamma^0 + \gamma^1 + \gamma^2 + \cdots + \gamma^{i-1})c^2 = (\frac{1}{2})^{2i}(\gamma^i - 1)/(\gamma - 1)c^2.$$

We can similarly find $var(i-1)$, and obtain the ratio between $var(i-1)$ and $var(i)$ to be

$$\frac{var(i-1)}{var(i)} = 4\frac{\gamma^{i-1} - 1}{\gamma^i - 1}.$$

Since $\gamma > 1$, for reasonably large stage number i, we thus have

$$\frac{var(i-1)}{var(i)} \approx \frac{4}{\gamma}.$$

Equating

$$\frac{4}{\gamma} = 2^{2H}$$

then completes our proof.

To better appreciate how the stage-dependent cascade model generates $1/f$ processes, we choose the stage-dependent PDFs to be Gaussian, as described by Eq. (9.18), and vary γ from 1.1 to 3.0 to obtain a number of random walk processes. Four examples for $\gamma = 1.4$, 1.6, 2.0, and 2.5 are shown in Fig. 10.2, where the variance-time plots are depicted. We observe excellent scaling laws in all four cases. The estimated Hurst parameters for these and other processes with different γ parameters are shown in Fig. 10.3 as triangles. Also shown in Fig. 10.3 is the variation of H against γ based on Eq. (10.2). We observe that Eq. (10.2) accurately estimates the Hurst parameter. Note that different estimators for the Hurst parameter give the same results.

A word on the numerical simulations of such processes is in order. One should always guarantee that

$$0 < \sigma^2(i, i+1) < 1/4$$

for all i. Again, this means that i cannot go to infinity, as we pointed out at the end of the previous section. In relation to fractal scaling law, this means that a power-law scaling relation cannot be valid for an infinite scaling regime. Indeed, physically observed fractal scaling relations are always truncated.

Our next theorem states that any $1/f$ process can be represented by such a cascade model.

Theorem 2 A $1/f^{(2H+1)}$ process can be represented as a stage-dependent random cascade process with $\gamma = 2^{2-2H}$.

Proof Suppose that the $1/f^{(2H+1)}$ process is represented by a time series y_n, $n = 1, 2, \cdots$. First, we obtain the increment process $x_n = y_n - y_{n-1}$. Then we add to the x_n time series a positive value $level$ so that the time series $\{x_n + level\}$ is positive. Denote this time series as X_n. For simplicity, assume that n runs from 1 to 2^N. We put X_j, $j = 1, \cdots, 2^N$, at stage N, according to our stage-dependent random cascade model. The weights at stages $N-1$, $N-2$, etc., can then be obtained by forming nonoverlapping running summations of length 2, 4, etc., as shown schematically in Fig. 10.4. Let two weights at stage $i-1$ and stage i be $X^{(i-1)}$ and $X^{(i)}$, respectively. Let $EX^{(i)} = ave(i)$ and $varX^{(i)} = \sigma^2(i)$. Then, $EX^{(i-1)} = 2ave(i)$, $varX^{(i-1)} = 2^{2H}\sigma^2(i)$. Now suppose that $X^{(i-1)}$ and $X^{(i)}$ are connected by the following simple relation,

$$X^{(i)} = r(i-1 \to i)X^{(i-1)},$$

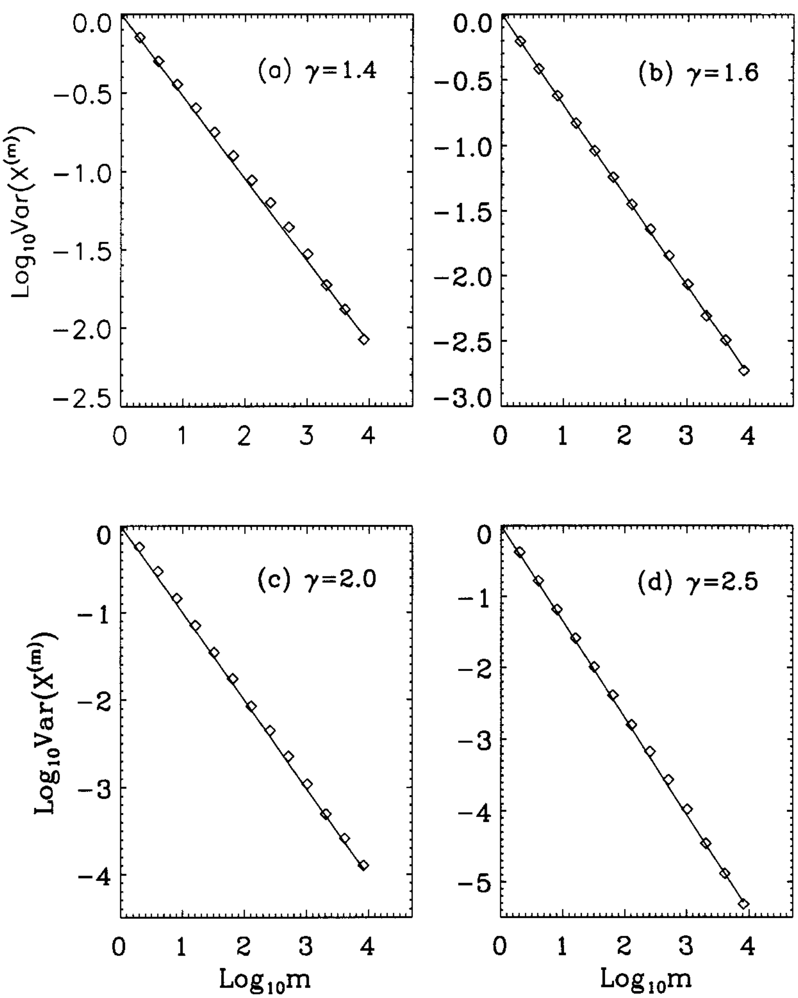

Figure 10.2. Variance-time plots for the time series generated by the stage-dependent cascade model with $\gamma = 1.4$, 1.6, 2.0, and 2.5.

where by construction, $r(i-1 \to i)$ and $X^{(i-1)}$ are independent. Then we have

$$ave^2(i) + \sigma^2(i) = [4ave^2(i) + 2^{2H}\sigma^2(i)][\frac{1}{4} + \sigma^2_{r(i-1\to i)}].$$

Hence,

$$\sigma^2_{r(i-1\to i)} = \frac{[1-2^{2H-2}]\sigma^2(i)}{4[ave^2(i) + 2^{2H-2}\sigma^2(i)]}.$$

Therefore,

$$\frac{\sigma^2_{r(i\to i+1)}}{\sigma^2_{r(i-1\to i)}} = \frac{2^{-2H}ave^2(i) + 2^{-2}\sigma^2(i)}{2^{-2}ave^2(i) + 2^{-2}\sigma^2(i)}. \tag{10.6}$$

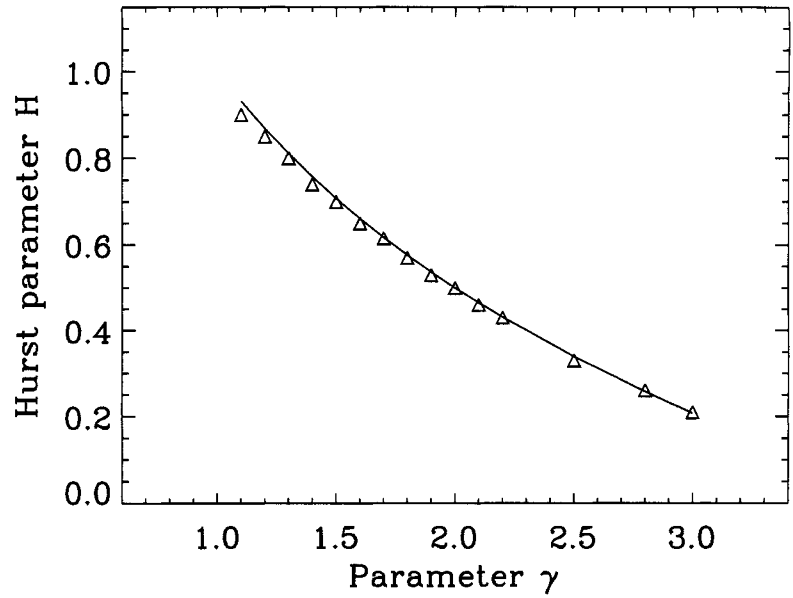

Figure 10.3. Variation of H vs. γ. The smooth curve is computed based on Eq. (10.2). Triangles are calculated based on variance-time plots such as shown in Fig. 10.2.

Stage										Scale ε
$\vdots$										$\vdots$
$N-3$	$X_1+X_2+X_3+X_4+X_5+X_6+X_7+X_8$								$\cdots$	$2^{-(N-3)}$
$N-2$	$X_1+X_2+X_3+X_4$				$X_5+X_6+X_7+X_8$				$\cdots$	$2^{-(N-2)}$
$N-1$	X_1+X_2		X_3+X_4		X_5+X_6		X_7+X_8		$\cdots$	$2^{-(N-1)}$
N	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	$\cdots$	2^{-N}

Figure 10.4. Schematic for obtaining weights at stages $N-1$, $N-2$, etc.

We can select *level* to be large enough so that $\sigma^2(i)$ is small compared to $ave^2(i)$. Hence

$$\frac{\sigma^2_{r(i\to i+1)}}{\sigma^2_{r(i-1\to i)}} = 2^{2-2H}.$$

Equating $\gamma = 2^{2-2H}$ then completes our proof.

As an example, let us consider a sequence of independent RVs *with exponential* distribution, $P(x) = \lambda e^{-\lambda x}$, $x \geq 0$. Such a time series has a Hurst parameter

$H = 1/2$. Since this is already a positive time series, we simply set our $level = 0$. At stage N, the mean value for the weights is $1/\lambda$, while the variance is $1/\lambda^2$. At stage i, $ave(i) = 2^{N-i}/\lambda$, $\sigma^2(i) = 2^{N-i}/\lambda^2$. Eq. (10.6) then becomes

$$\frac{\sigma^2_{r(i\to i+1)}}{\sigma^2_{r(i-1\to i)}} = \frac{2 \cdot 2^{N-i} + 1}{2^{N-i} + 1} \approx 2 \tag{10.7}$$

for reasonably large $N - i$ values.

Actually, for the above example, the distribution for $r(i \to i+1)$ can be readily obtained. At the $(N - i)$th stage of the cascade process, the random variables, $w_j^{(N-i)}$, $1 \le j \le 2^{(N-i)}$, are sums of 2^i iid exponential random variables. Hence, $w_j^{(N-i)}$'s are iid 2^i-Erlang random variables with mean $\frac{2^i}{\lambda}$ and variance $\frac{2^i}{\lambda^2}$. The density function of an m-Erlang random variable (i.e., the sum of m iid exponential random variables), X, which is a special case of the gamma random variables, is given as

$$f_X(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{m-1}}{(m-1)!},$$

where $1/\lambda$ is the mean of the individual exponential RVs. For the sake of notational convenience, let us represent two consecutive odd- and even-numbered RVs at the $(N - i)$th level as X and Y, respectively. Then, the corresponding weight at the $(N - i - 1)$th stage is given by $Z = X + Y$ and the multiplier $r((N - i) - 1 \to (N-i)) = \frac{X}{X+Y}$. Clearly, Z is now a $2^{(i+1)}$-Erlang random variable with twice the mean and the variance of X. Fortunately, the multiplier $r((N - i - 1) \to (N - i))$ also belongs to a well-studied class of random variables, called the beta random variables. The density function of $r((N - i - 1) \to (N - i))$ is given as

$$f_{r((N-i-1)\to(N-i))}(r) = \frac{(2m-2)!}{((m-1)!)^2}(r(1-r))^{m-1}, \tag{10.8}$$

where $m = 2^i$. Clearly, $r((N - i - 1) \to (N - i)$ is symmetric around $1/2$, and its variance is given by

$$\sigma^2((N - i - 1) \to (N - i)) = \frac{1}{4(2^{i+1} + 1)}.$$

Thus, the multiplicative parameter $\gamma_{(N-i)}$ (for $i \ge 1$) is given by

$$\gamma_{(N-i)} = \frac{\sigma^2((N-i) \to (N-i+1))}{\sigma^2((N-i-1) \to (N-i))} = \frac{2^{i+1}+1}{2^i+1} \tag{10.9}$$

and is identical to Eq. (10.7).

Moreover, one can show that the random variables Z and $r((N-i-1) \to (N-i))$ are statistically independent. Hence, the stage-dependent cascade model can *exactly generate an SRD process comprising a sequence of independent RVs with identical exponential distributions*. In the model, choose the multiplier random variables connecting consecutive stages $(N-i)-1$ and $(N-i)$ $(0 \leq i \leq N-1)$ independently and identically from the distribution given by Eq. (10.8).

As another example, let us consider independent, uniformly distributed RVs in the unit interval. Here again we have $H = 1/2$. Set $level = 0$. At stage N, the mean and variance of the time series are 1/2 and 1/12, respectively. At stage i, $ave(i) = 2^{-1} \cdot 2^{N-i}$, $\sigma^2(i) = 2^{N-i}/12$. Thus, Eq. (10.6) becomes

$$\frac{\sigma^2_{r(i \to i+1)}}{\sigma^2_{r(i-1 \to i)}} = \frac{6 \cdot 2^{N-i} + 1}{3 \cdot 2^{N-i} + 1} \approx 2 \tag{10.10}$$

for reasonably large $N - i$ values.

To assess how well a cascade model represents a $1/f$ process with different H parameters, we numerically study fGn processes with $H = 0.75$ and 0.25, as well as a sequence of independent RVs with exponential distribution and uniform distribution. Following the procedures described in the proof, we compute $\sigma^2_{r(i \to i+1)}$ and plot $\log_2 \sigma^2_{r(i \to i+1)}$ vs. the stage number. The results are shown in Fig. 10.5. We observe excellent scaling relations in all these cases, with the slopes of those straight lines given by $2 - 2H$, thus γ given by Eq. (10.2).

10.3 APPLICATION: MODELING HETEROGENEOUS INTERNET TRAFFIC

10.3.1 General considerations

For simplicity, in this section we only consider the modeling of counting traffic processes. The discussion will, of course, also be pertinent to the modeling of any other interesting time series.

Suppose that we have a positive time series of length 2^N. We want to model it using our stage-dependent model. For simplicity, we may choose all the PDFs to be of the same functional form, say the truncated Gaussian (Eq. (9.18)). In the ideal setting of the stage-dependent cascade model, we constrain the model by Eq. (10.1). So the model contains two independent parameters, γ and $\sigma^2(0,1)$ or $\sigma^2(N-1,N)$. What are the physical meanings of these two parameters? Following the arguments in Sec. 10.1, we can readily understand that they are the burstiness indicators: when γ is fixed, the burstiness of the modeled traffic increases when $\sigma^2(0,1)$ or $\sigma^2(N-1,N)$ is increased. When the latter is fixed, the burstiness of the modeled traffic decreases when γ is increased. Since γ is related to the Hurst parameter by Eq.(10.2), it is clear that the Hurst parameter alone is not the burstiness indicator. More interestingly, this model generates an ideal multiplicative multifractal ($\gamma = 1$), conventional LRD processes ($1 < \gamma < 2$), and short-range-dependent processes such as Poisson processes ($\gamma = 2$) and antipersistent processes ($2 < \gamma < 3$) as special cases. Equally important is that this process makes us realize that the boundary between conventional LRD traffic models and Markovian traffic models is far more vague than one might have thought.

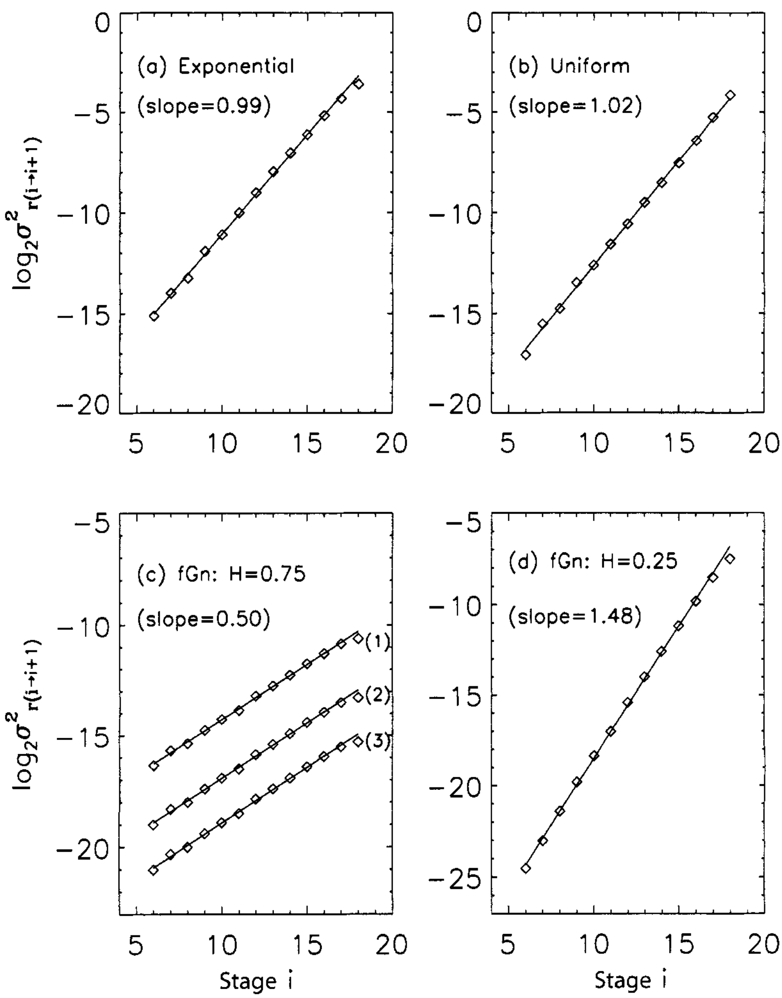

Figure 10.5. $\log_2 \sigma^2_{r(i\to i+1)}$ vs. the stage number for a sequence of independent RVs with exponential distribution (a) and with uniform distribution (b), and for fGn processes with $H = 0.75$ (c) and $H = 0.25$ (d). The three lines in (c) correspond to three different choices of $level$. We see that $level$ does not affect the slope of these straight lines.

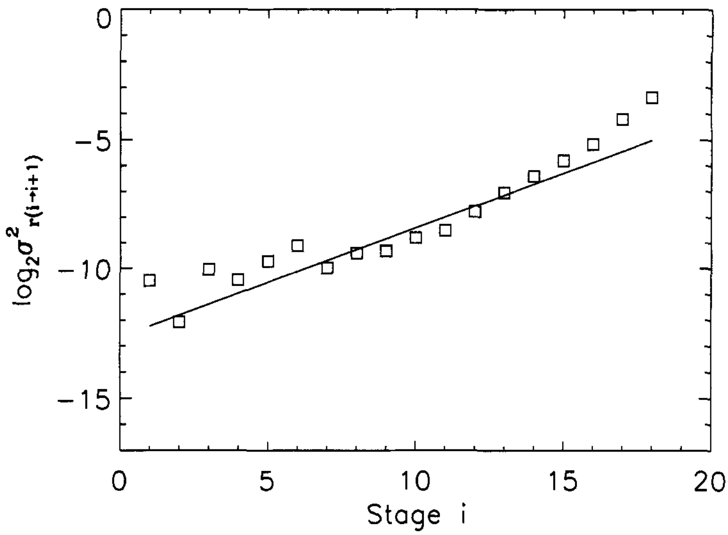

Figure 10.6. $\log_2 \sigma^2_{r(i\to i+1)}$ vs. stage i for the counting process of OSU (square symbols). The straight line is a least linear squares fit of the square symbols.

There are two specific important advantages of such modeling over the fBm processes: these processes have a non-Gaussian nature and are easier to simulate.

At the next level of complexity, one may wish the model to be constrained by two conditions similar to Eq. (10.1). That is to say, the modeled traffic has two distinct fractal scaling laws. We need four parameters to specify the model — γ_1, γ_2, $\sigma^2(0,1)$, or $\sigma^2(N-1,N)$ — and a parameter specifying where the transition from one type of fractal to another occurs. When one of the γ is one, then the number of parameters in the model is reduced to three, the model is an ideal multiplicative multifractal in one scaling region and a $1/f$ process in another scaling region. Following this line of argument, it is clear that, under the constraint that all PDFs $P^{(i,i+1)}(r)$ are truncated Gaussian (Eq. (9.18)), the most general stage-dependent cascade model will be specified by N variances, $\sigma^2(i-1,i)$, $i = 0,\cdots,N-1$. It is this type of model (in the wavelet domain) that has been used by Feldmann et al. [134], Gilbert et al. [187], and Riedi et al. [368]. It should now be clear why their models are not truly multifractals. Nevertheless, Feldmann et al. still bravely went on to explain the mechanism for their model based on the layered structure of a network. We should emphasize here that any random function may be approximated by an unconstrained stage-dependent model with any precision.

10.3.2 An example

As an example, we model the very-high-speed Backbone Network Service (vBNS) traffic trace OSU (see Sec. A.1 of Appendix A for the description of the trace data). We analyze the data using the same procedure for the analysis of the data presented

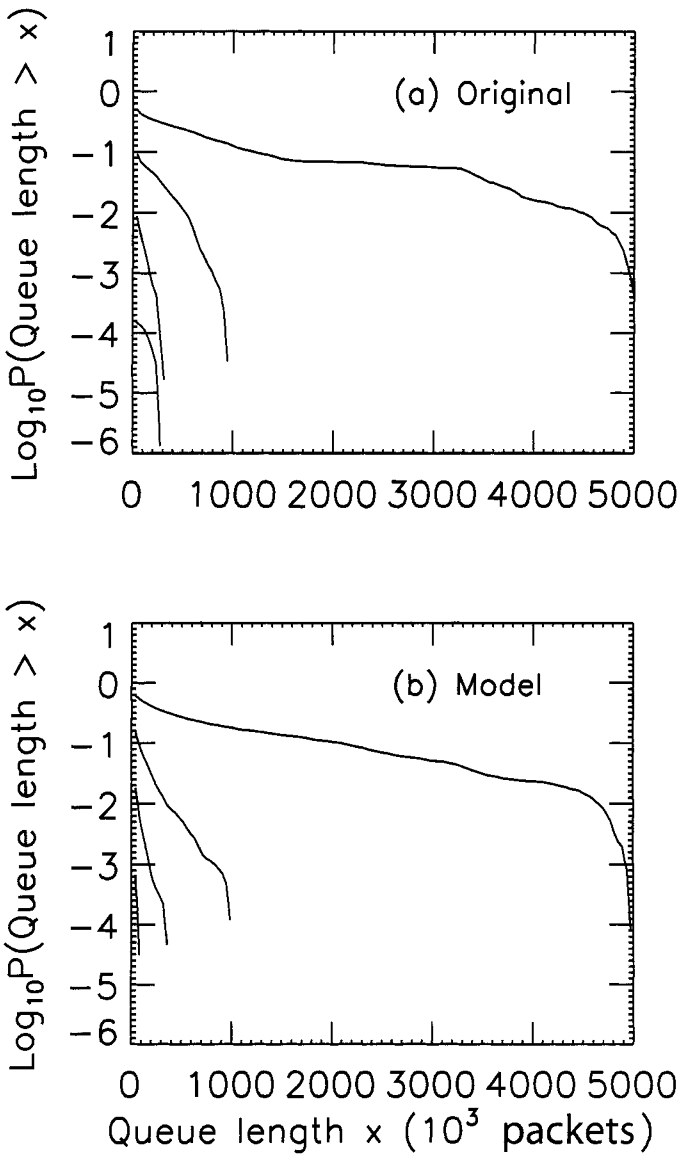

Figure 10.7. System size tail probabilities obtained when the traffic trace OSU (a) and modeled stage-dependent multiplicative process traffic (b) are used to drive the queuing system. Four curves, from top to bottom, correspond to $\rho = 0.9$, 0.7, 0.5, and 0.3, respectively.

in Fig. 10.5. The result is shown in Fig. 10.6 as the square symbols for the counting process of the OSU data. Note that when i is small, $\sigma^2(i, i+1)$ may not be well estimated due to the sparsity of data points. So we conclude that at least for $i > 7$, $\log_2 \sigma^2_{r(i \to i+1)}$ vs. i follows approximately a straight line. However, we start from a "global" (i.e., for all i) least linear squares fitting of those square symbols. The fitted line is shown in Fig. 10.6 as the solid line. Next, we generate a modeled counting traffic process according to that fitted line, and drive a FIFO single-server queuing system by the modeled traffic as well as by the original OSU traffic trace. The results are shown in Fig. 10.7. We note that for utilization levels $\rho = 0.9, 0.7$, and 0.5, such a model is already excellent. The deviation observed in Fig. 10.7 when $\rho = 0.3$ is due to the fact that we have actually fitted a straight line for all i in Fig. 10.6. Thus, if we are willing to increase the number of parameters slightly, the model will be more accurate.

10.4 BIBLIOGRAPHIC NOTES

The model discussed here was first proposed in [174]. For a proof on the independence between Z and $r((N - i - 1) \to (N - i))$, see pp. 60–61 of [375]. For traffic modeling, see [134, 187, 368].

10.5 EXERCISES

1. Reproduce Fig. 10.2.
2. Reproduce Fig. 10.3.
3. Reproduce Fig. 10.5.