Multiplicative Cascade Modeling of Computer Network Traffic

Patricia H. Carter

B10, NSWCCDD

Interface 2002

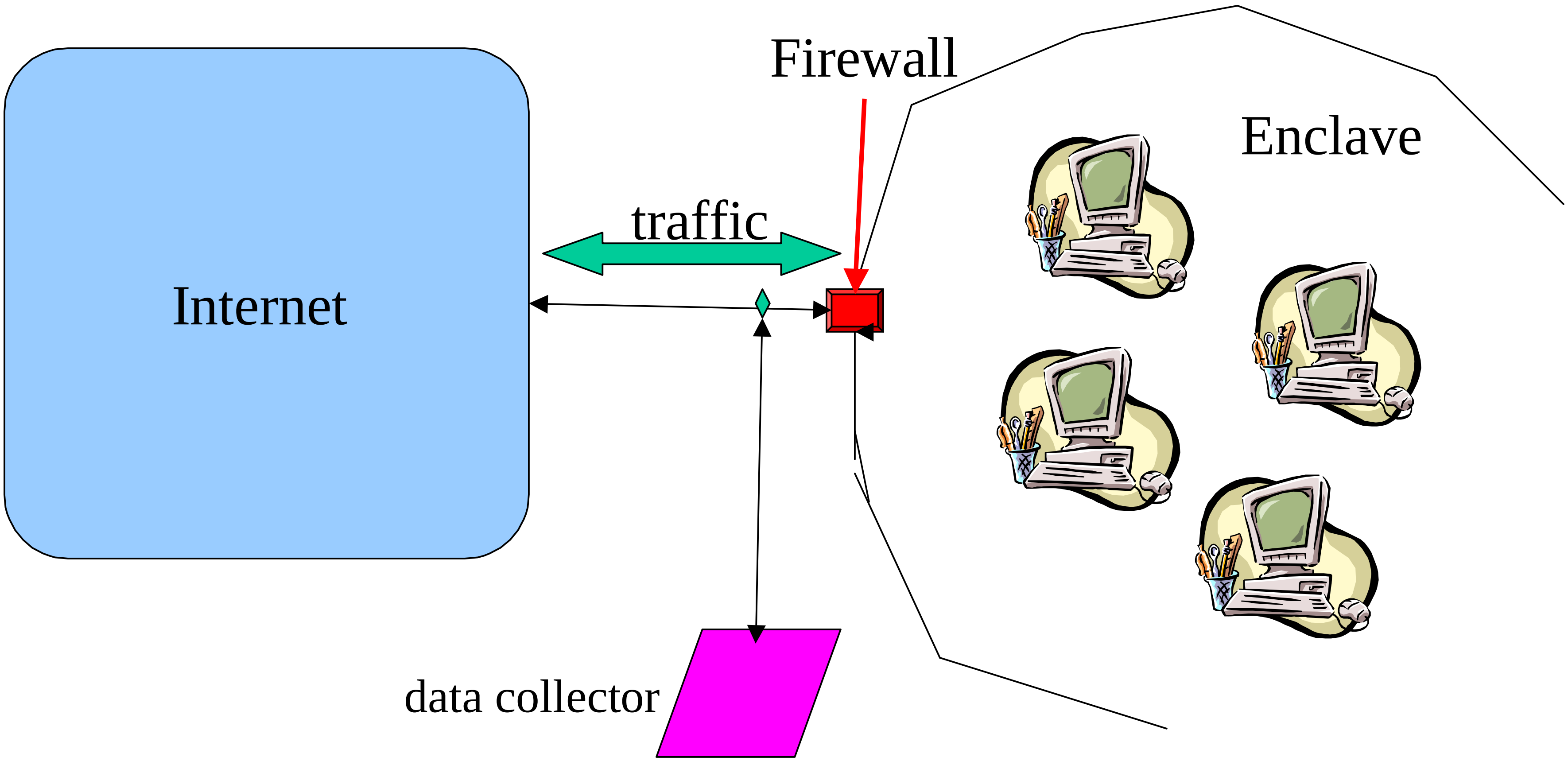
April 19,2002

outline

- network traffic data
- the multiplicative cascade
- visualizing the cascade
- measuring burstiness
    - the structure function
    - the multifractal spectrum
- conclusions

NAVSEA
DAHLGREN
Surface Warfare Center Division

Wide Area Network traffic collection at enclave boundary

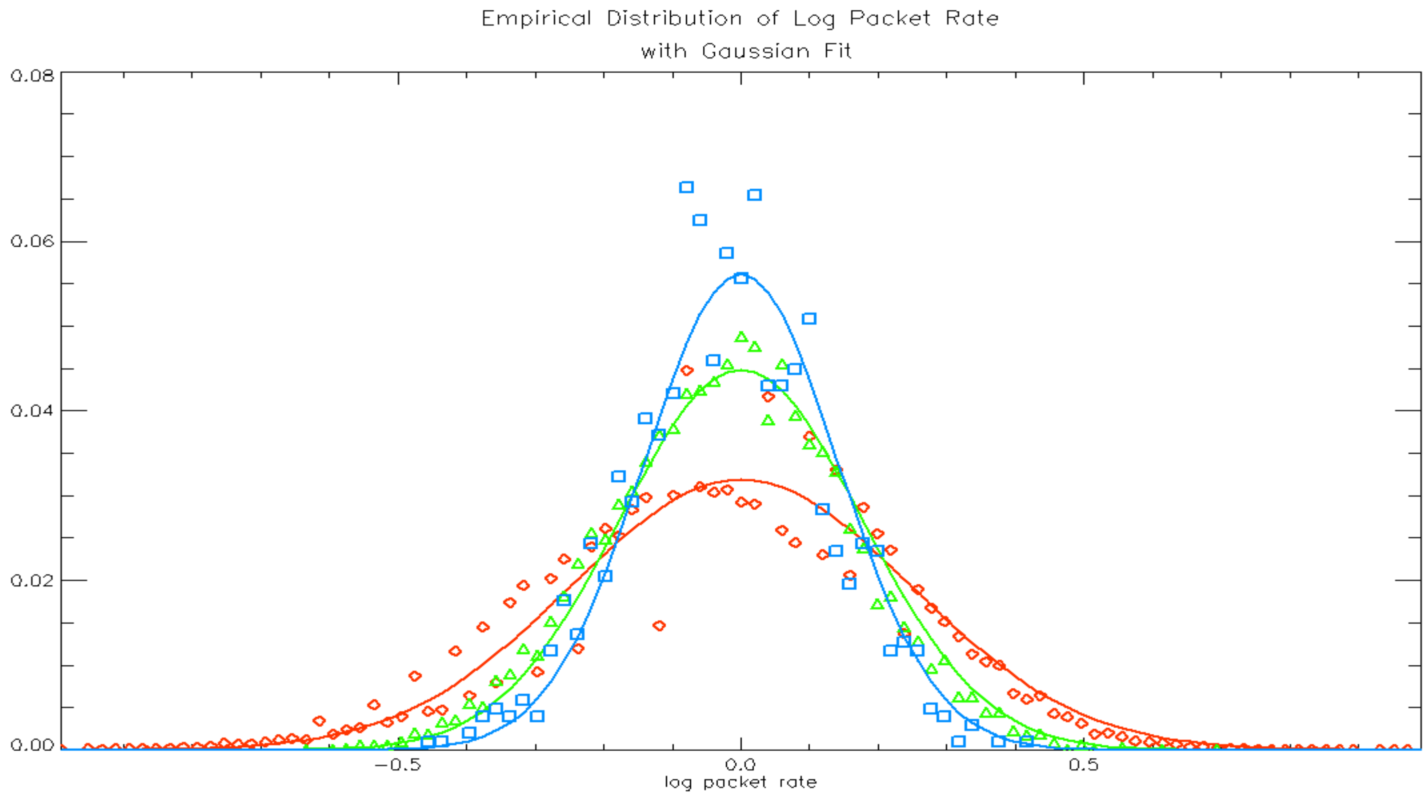


Packet Rate Process

- TCP packets entering/leaving protected network
- raw data is arrival times
- packet rate process is # of packets/unit time

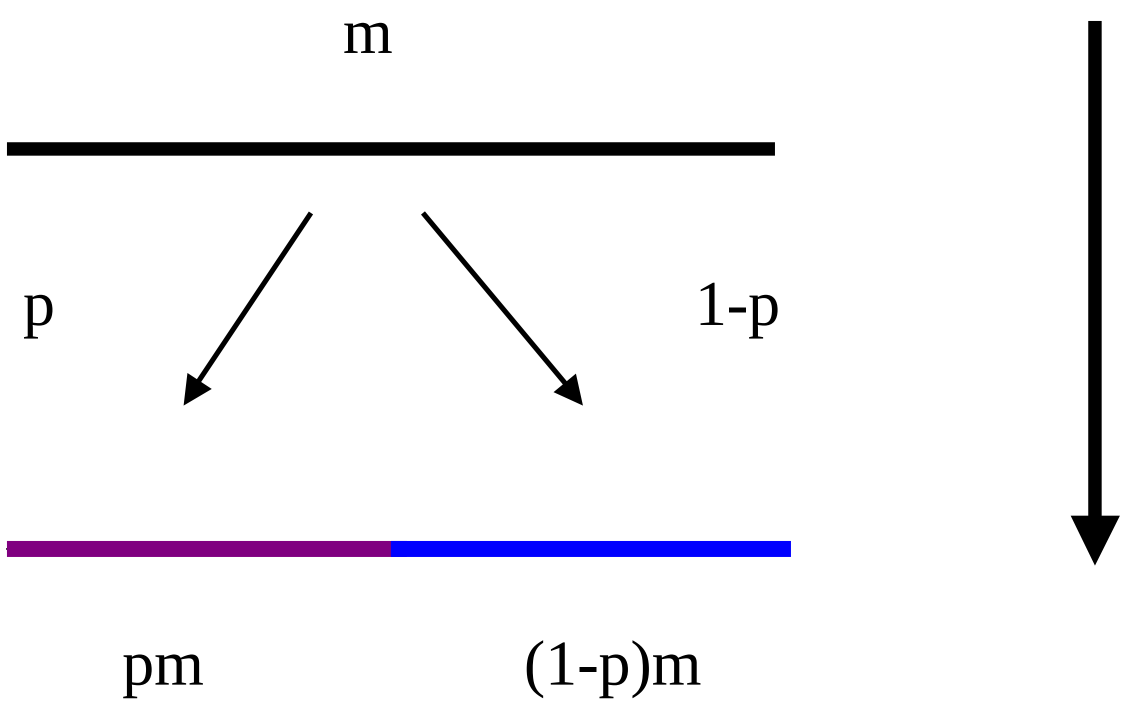
Three Resolutions of Packet Rate Data

Packet Rate Process "approximately" Log Normal – hour 12



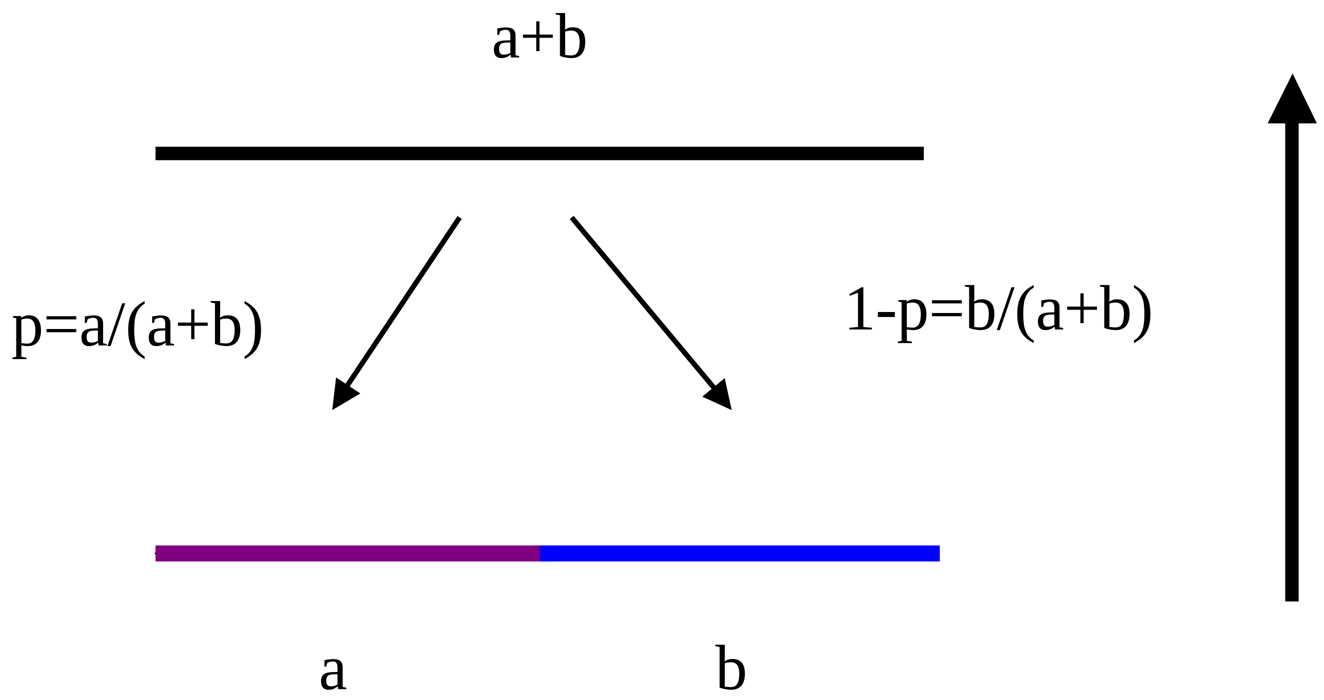
Packet rate distribution is approximated log normal.

Multiplicative Cascade: *synthesis*



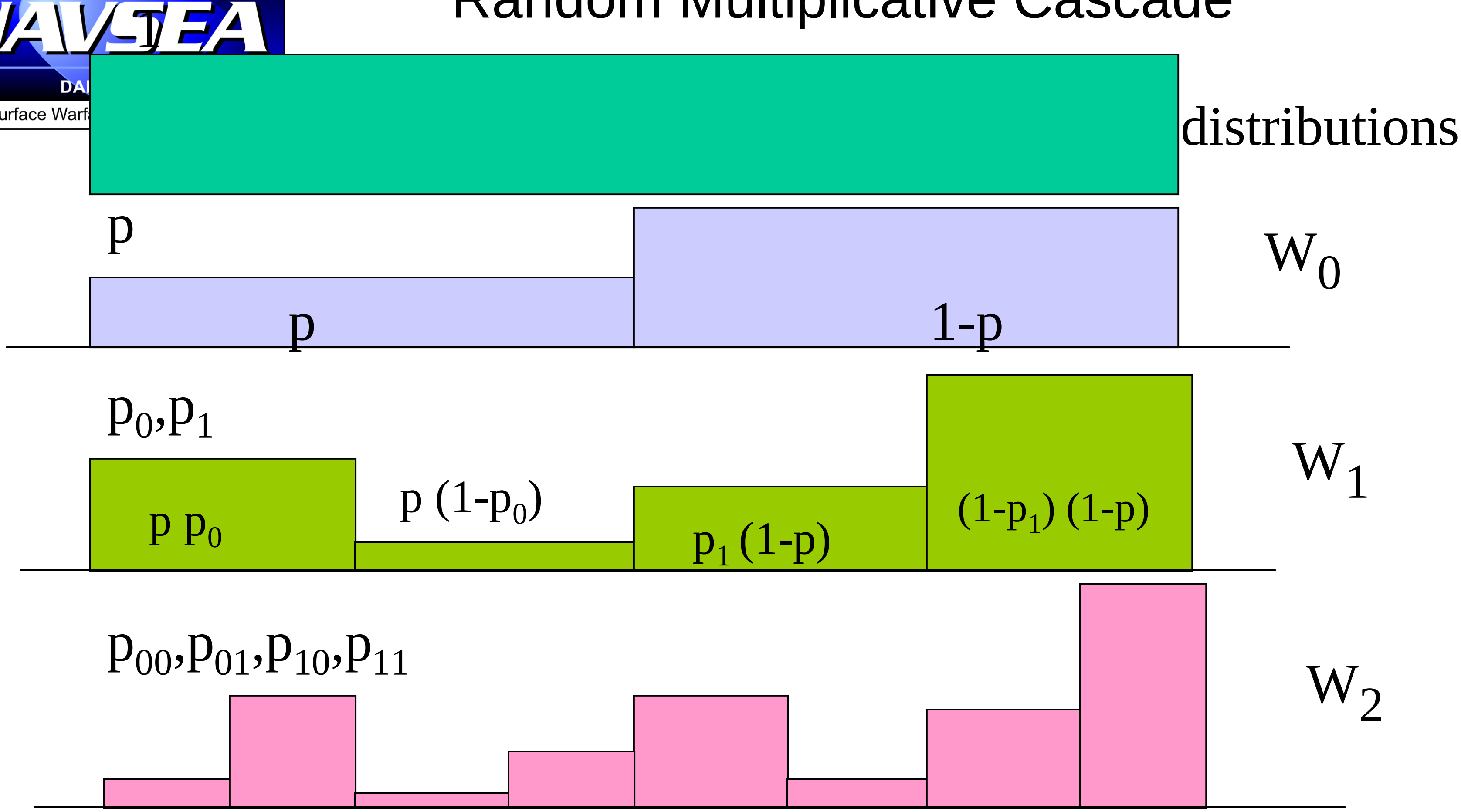
P is a random variable from a distribution supported on [0,1] with mean ½ and variance v – **conservative cascade**

Multiplicative Cascade:
analysis



If (a+b)=0 then choose p uniformly from {0,1} .

Random Multiplicative Cascade

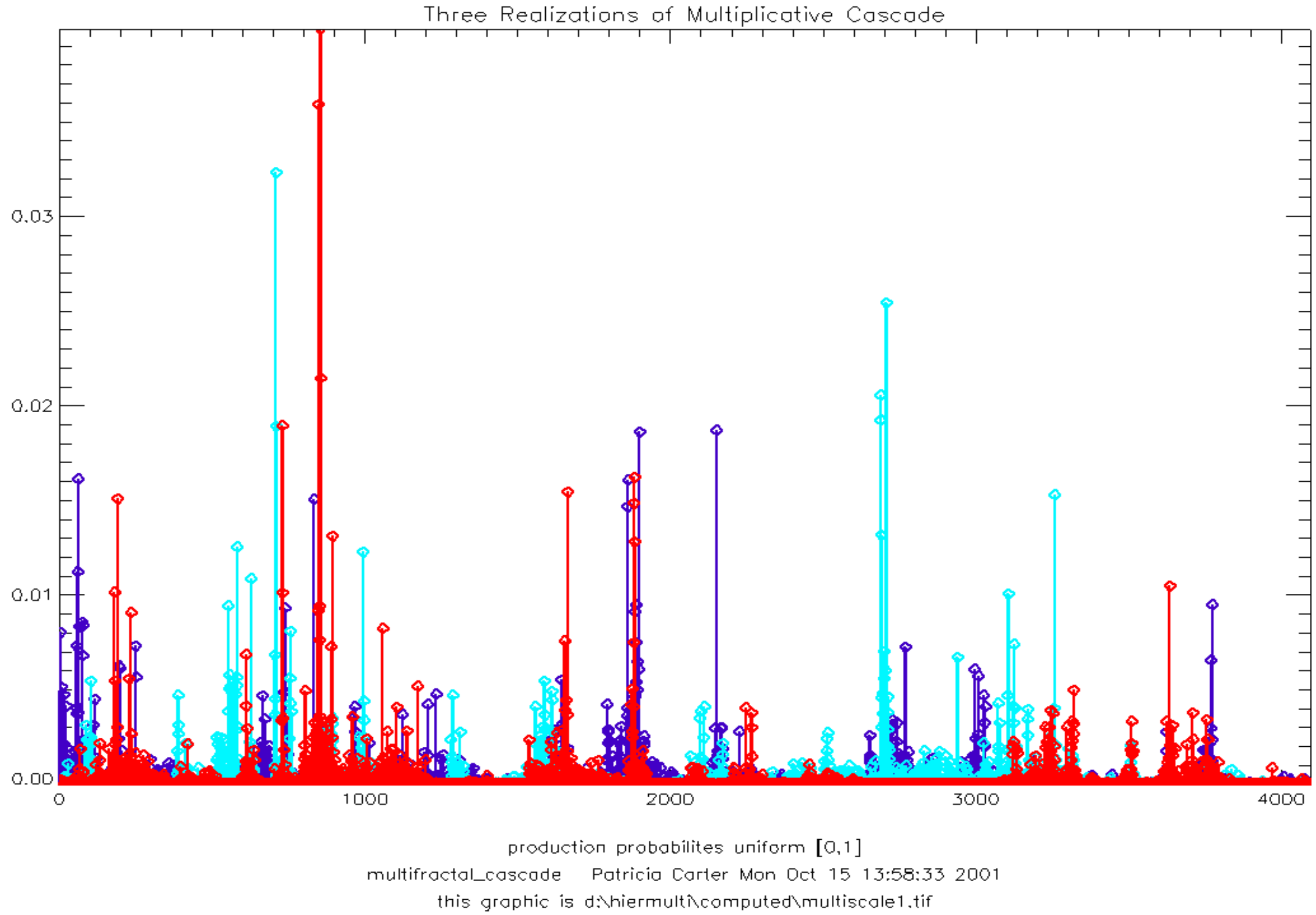


If the distributions are all the same, one example, chosen from the beta distribution whose density is

$$f(u) = u^a (1-u)^{a-1} \Gamma(2a) / \Gamma(a)^2$$

Random Multiplicative Cascade

Synthetic data – three realizations



the vector P of multipliers

Suppose packet rate process R has 2^L samples

next smallest scale p's

$$p_{2^{L-2}}, \ldots, p_{2^{L-1}-1}$$

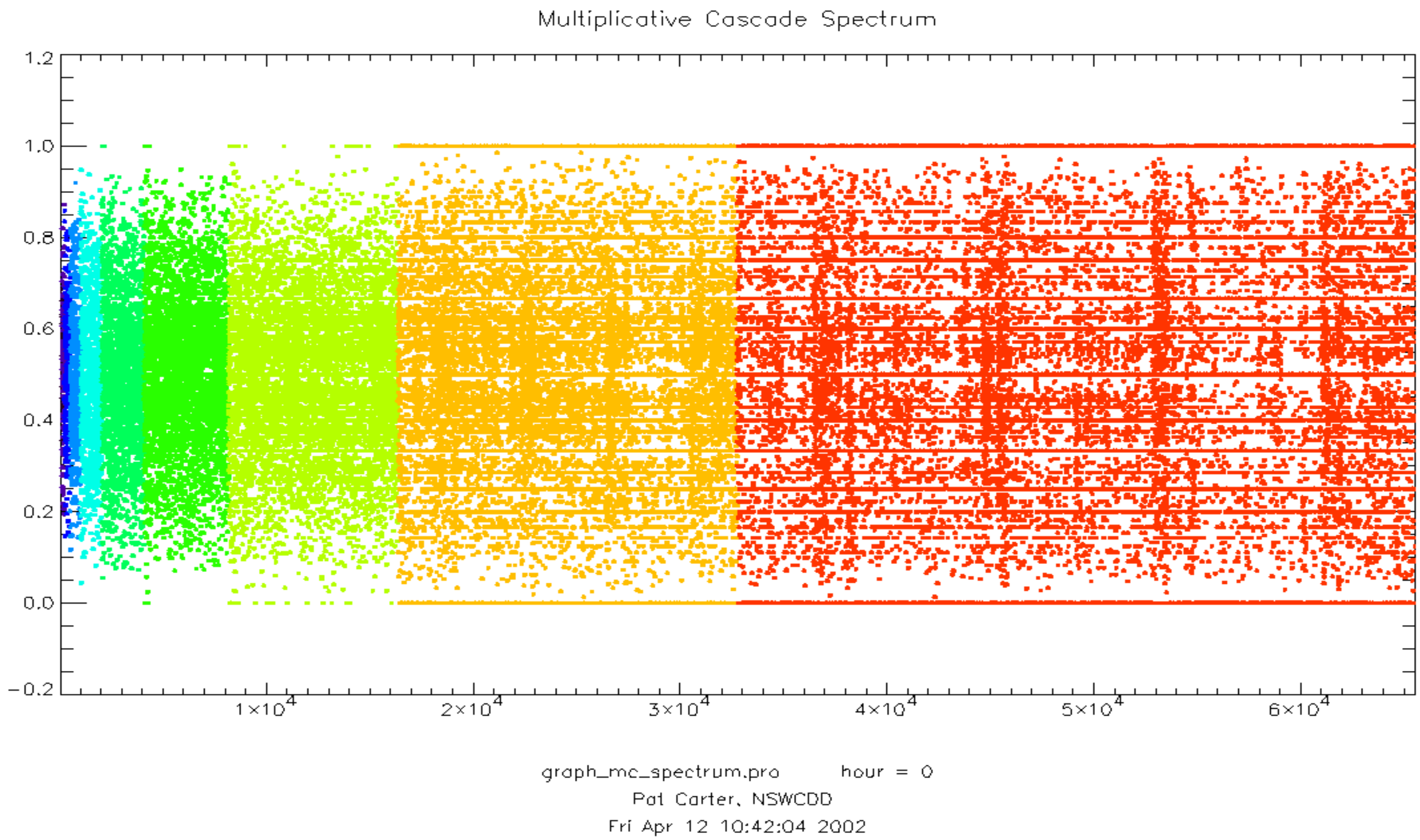
smallest scale p's in time order

$$p_{2^{L-1}}, \ldots, p_{2^L-1}$$

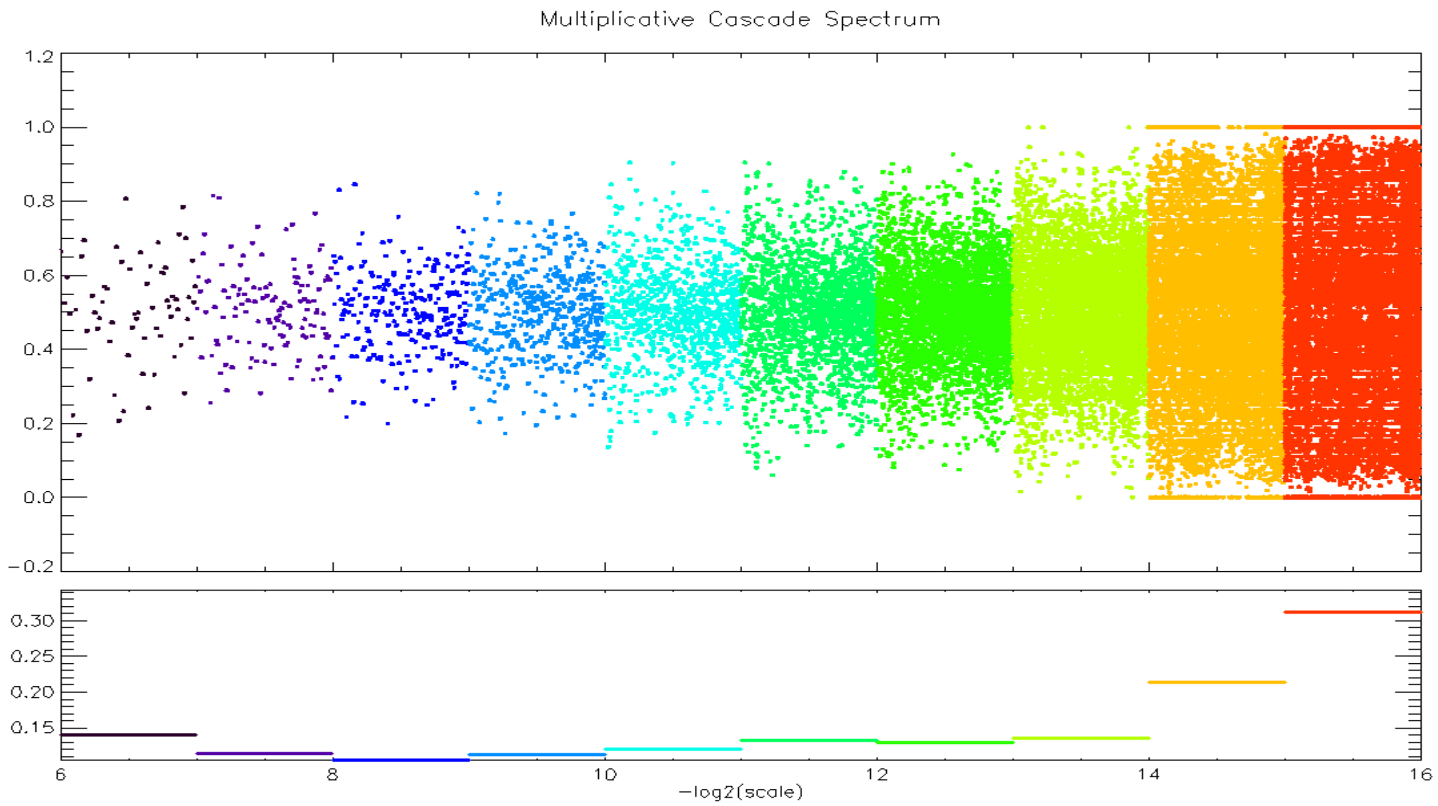
Finally $$P_0 = \sum_{i=0}^{2^L-1} R_i$$

So the P and R are a "transform pair".

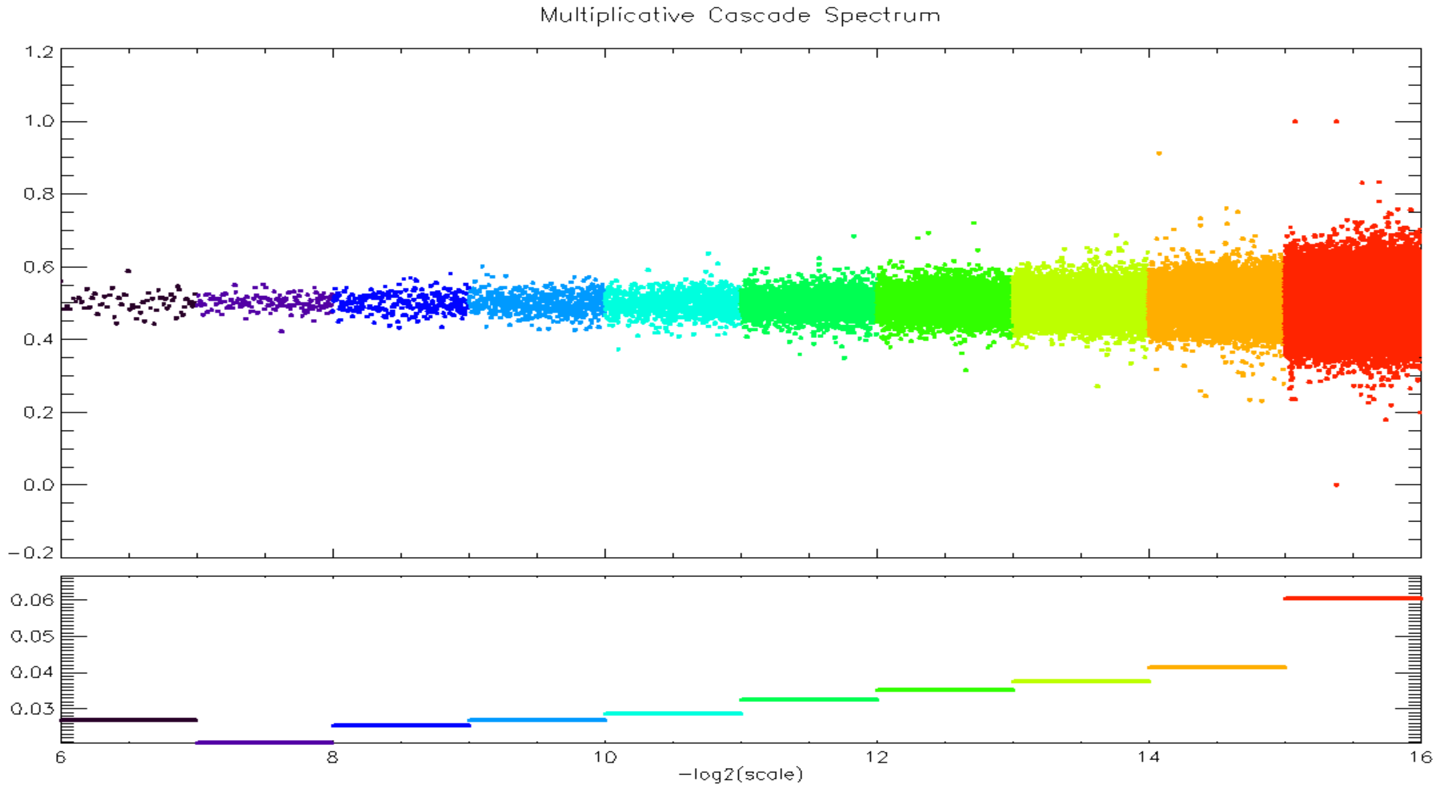
Multipliers calculated via inverse cascade procedure



Multipliers Plotted as a Function of Scale – hour 0

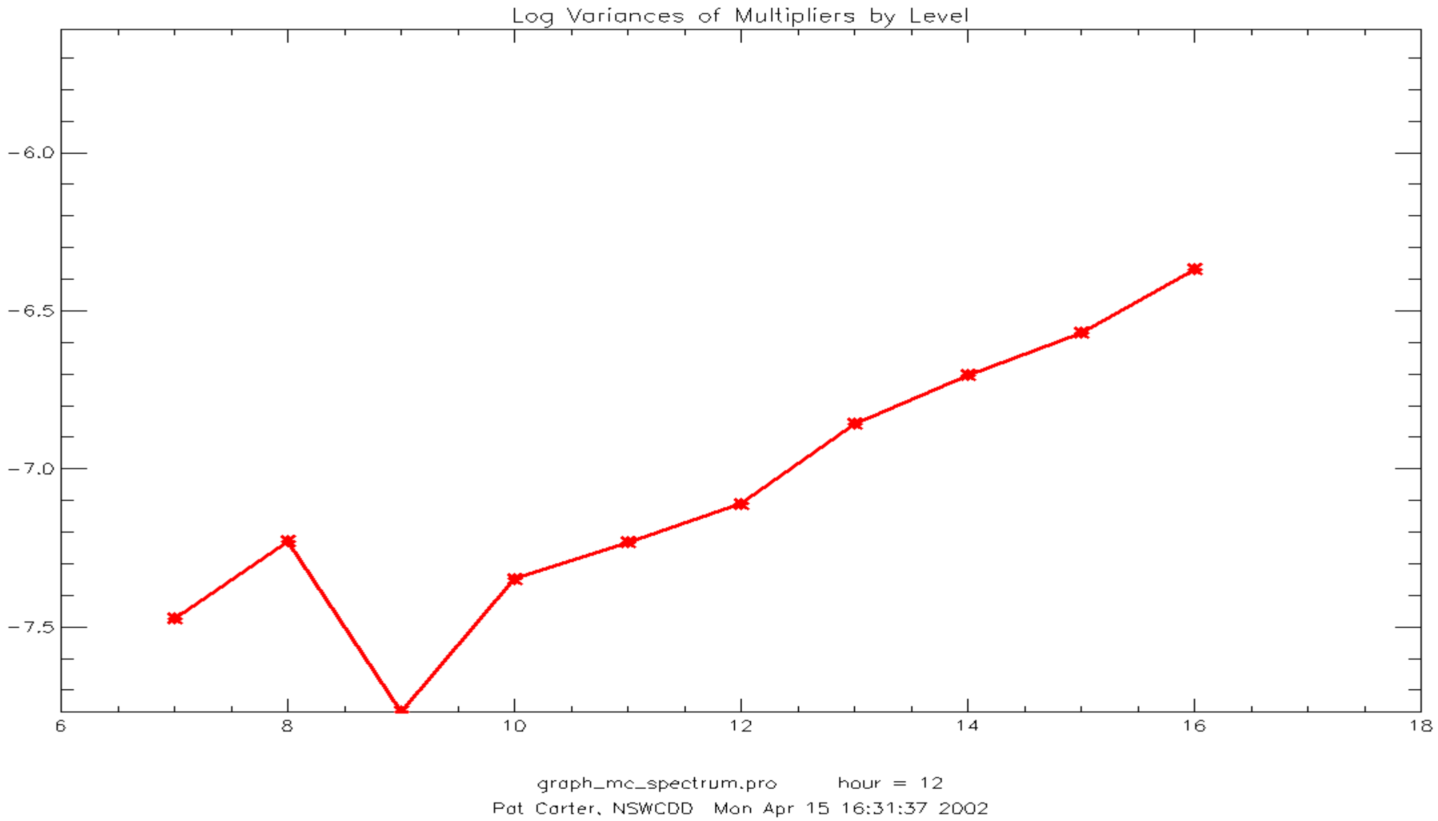


Multipliers Plotted as a Function of Scale – hour 12

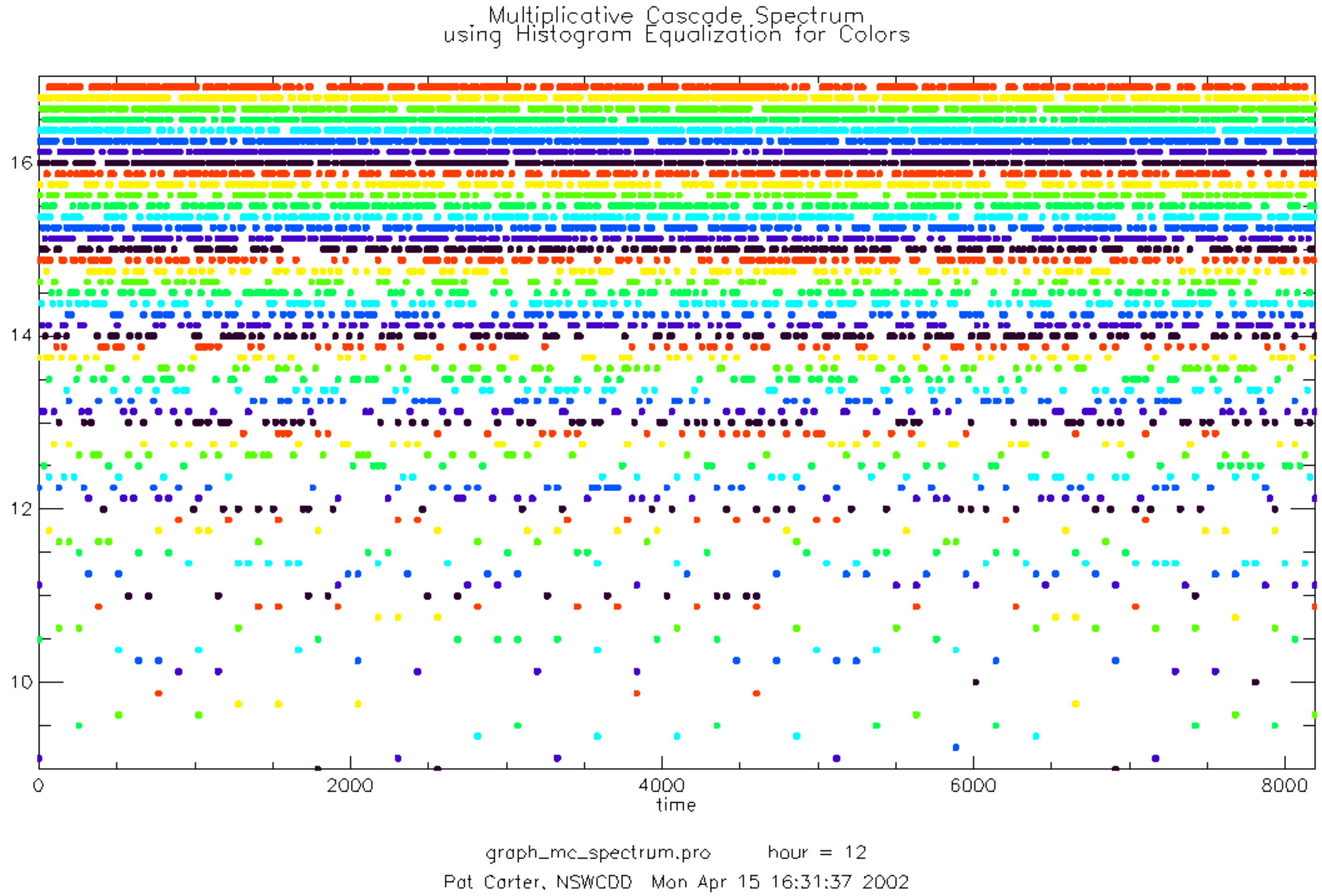


NAVSEA DAHLGREN
Surface Warfare Center Division

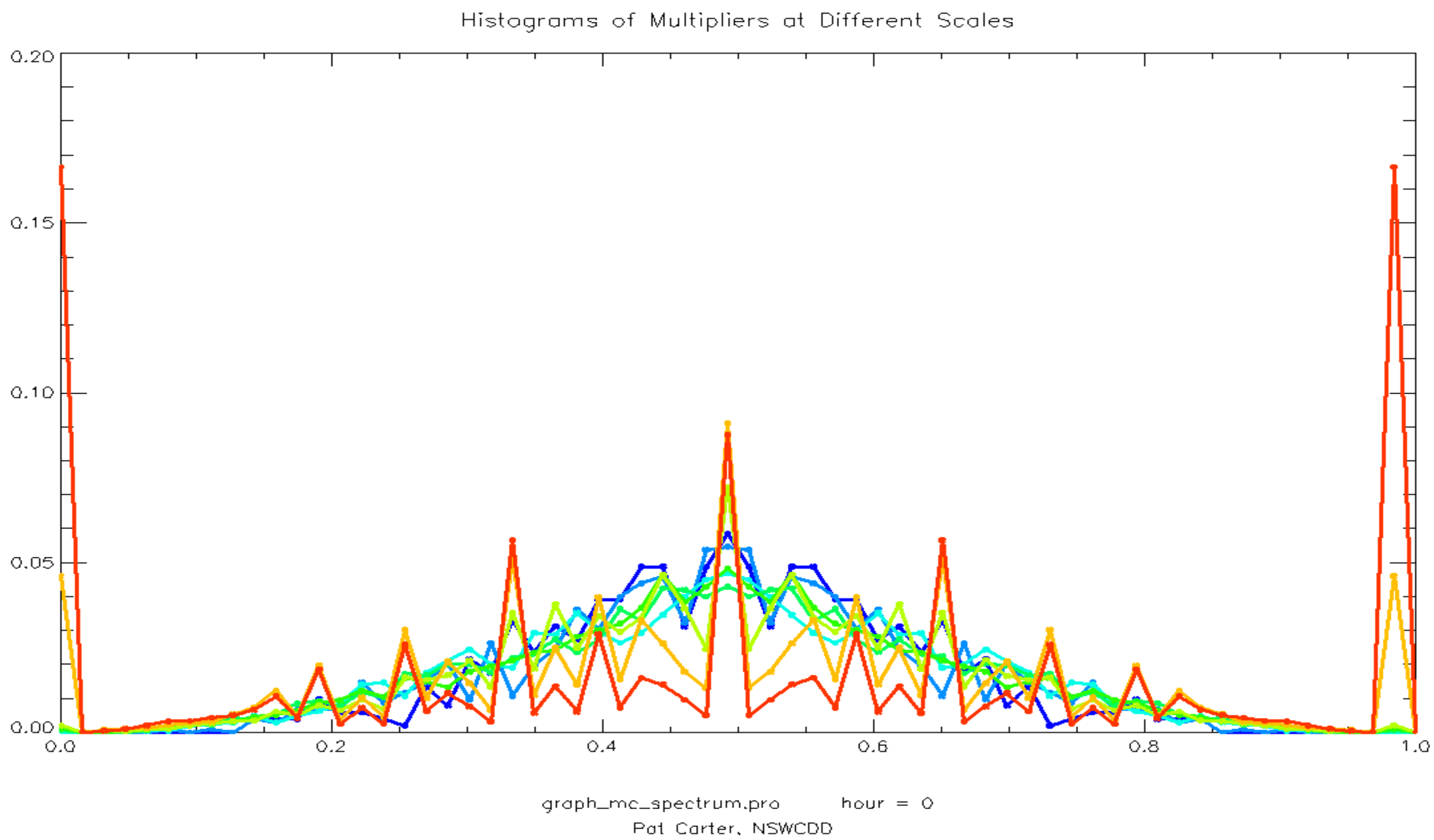
Log Variances of Multipliers versus Log Scale - Hour 12



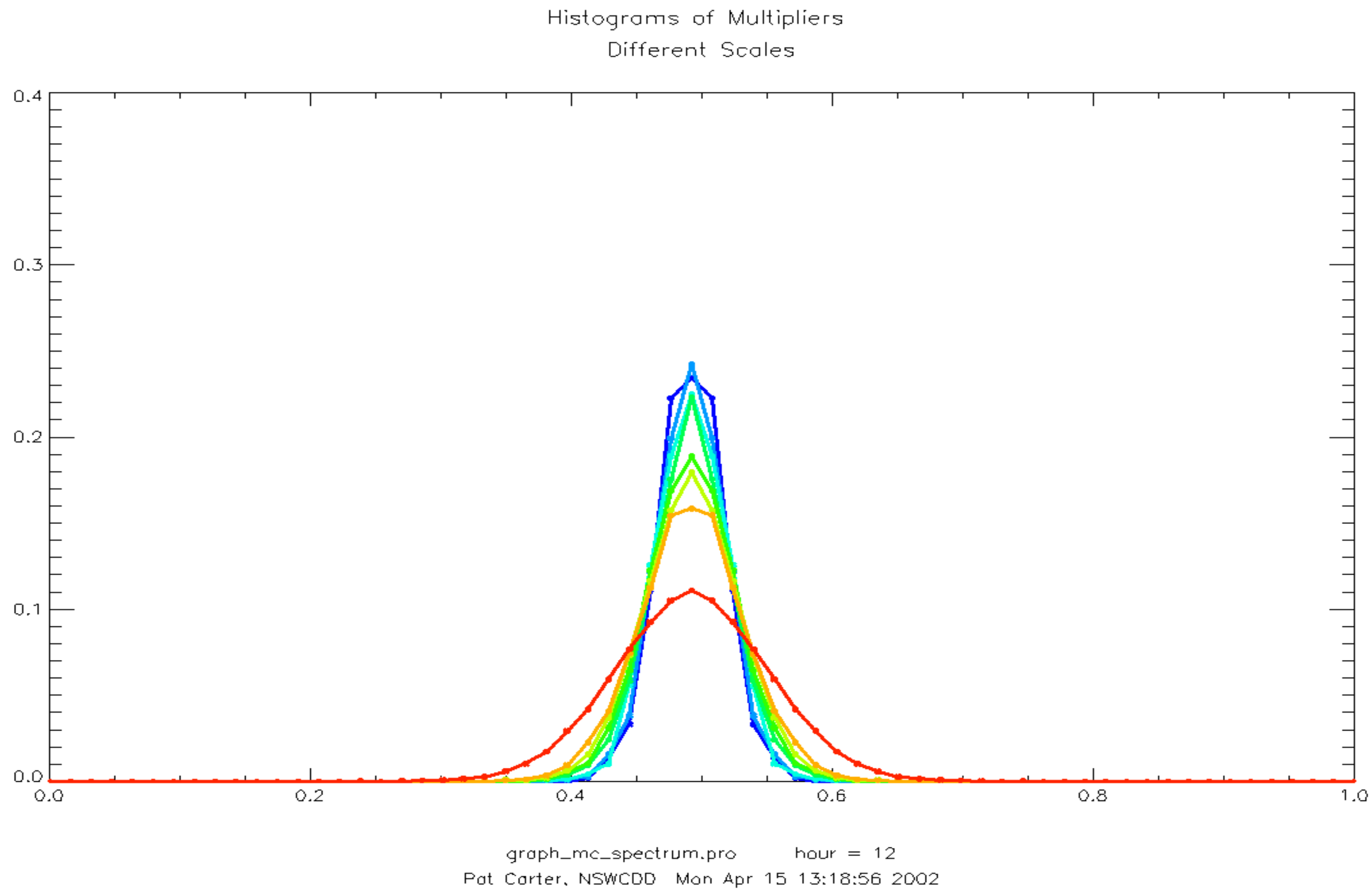
Time-Scale Visualization
Histogram-equalized



Histograms of Multipliers at Each Scale – Hour 0



Histograms of Multipliers at Each Scale – Hour 12



The Structure Function

$$T(q) = 1 + \log\left(\sum_{i=1}^{2^L-1} p_i^{\,q} / (2^L - 1)\right)$$

This assumes the multiplier distributions are the same at every scale, but they aren't.

NAVSEA DAHLGREN Surface Warfare Center Division

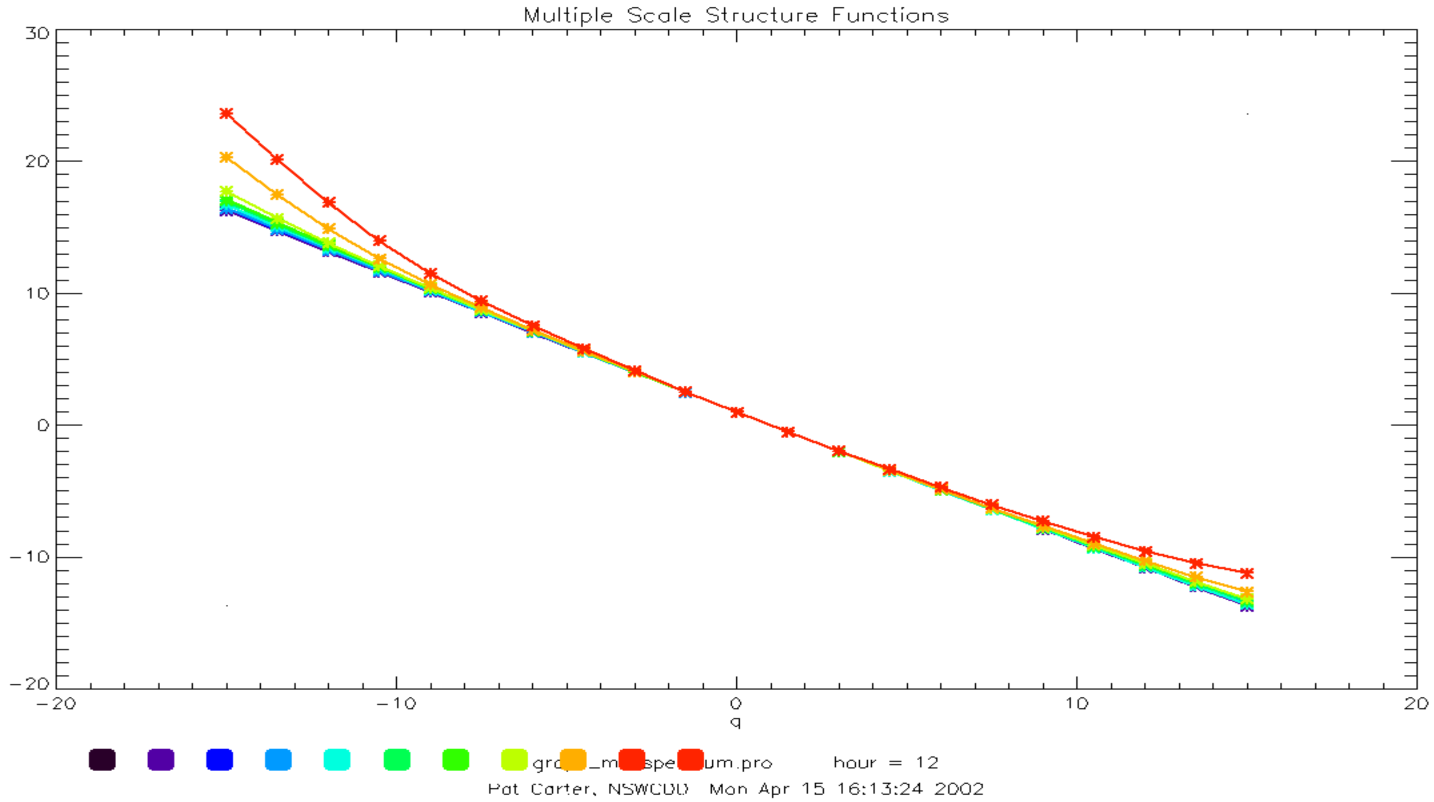
Multiple Scale Structure Function

$$T(q, M) = 1 + \log \sum_{i=2^{M-1}}^{2^M - 1} p_i^{\,q} / 2^{M-1}$$

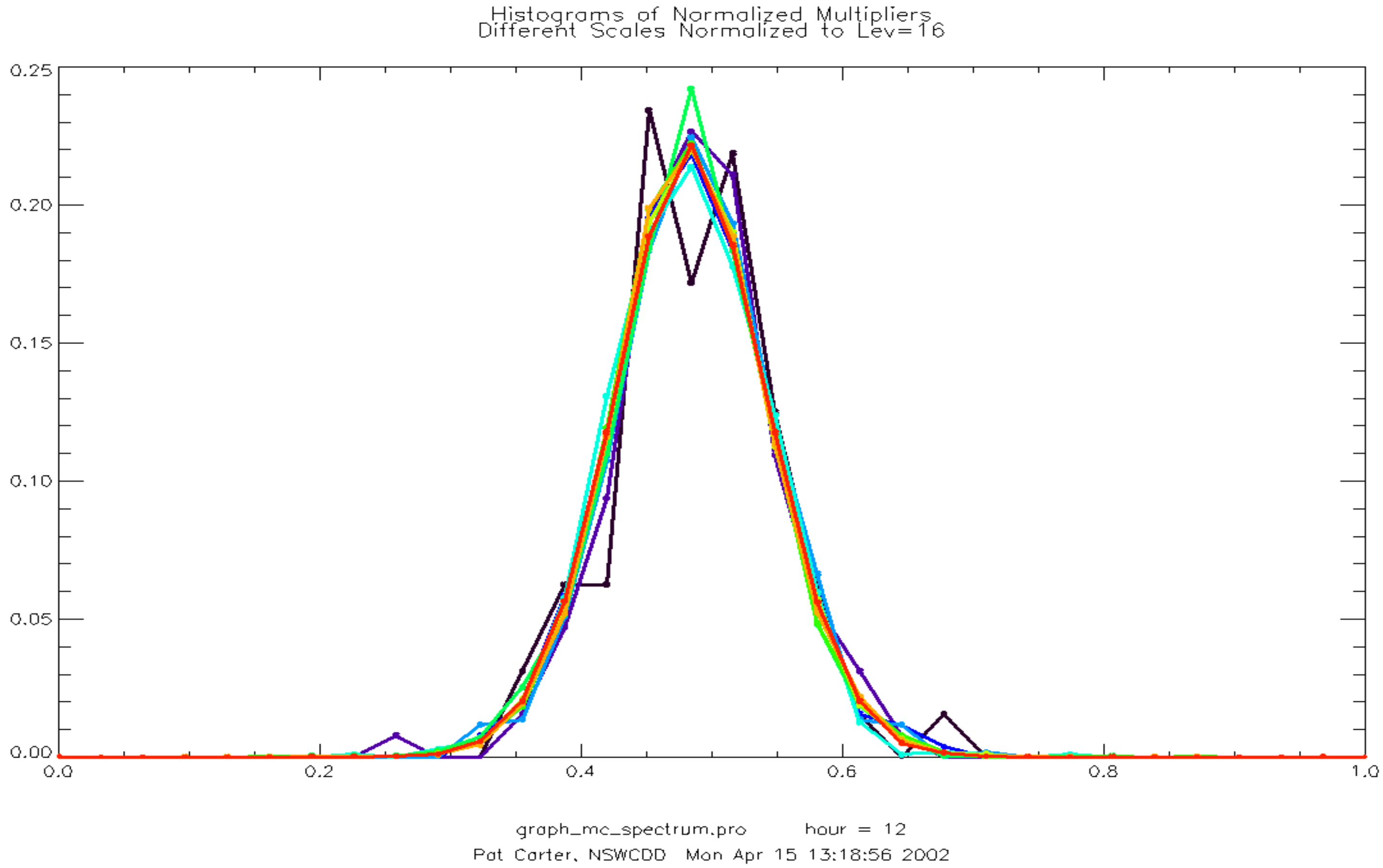
Where $\{p_i : i = 2^{M-1}, 2^M - 1\}$

are the multipliers calculated at level M.

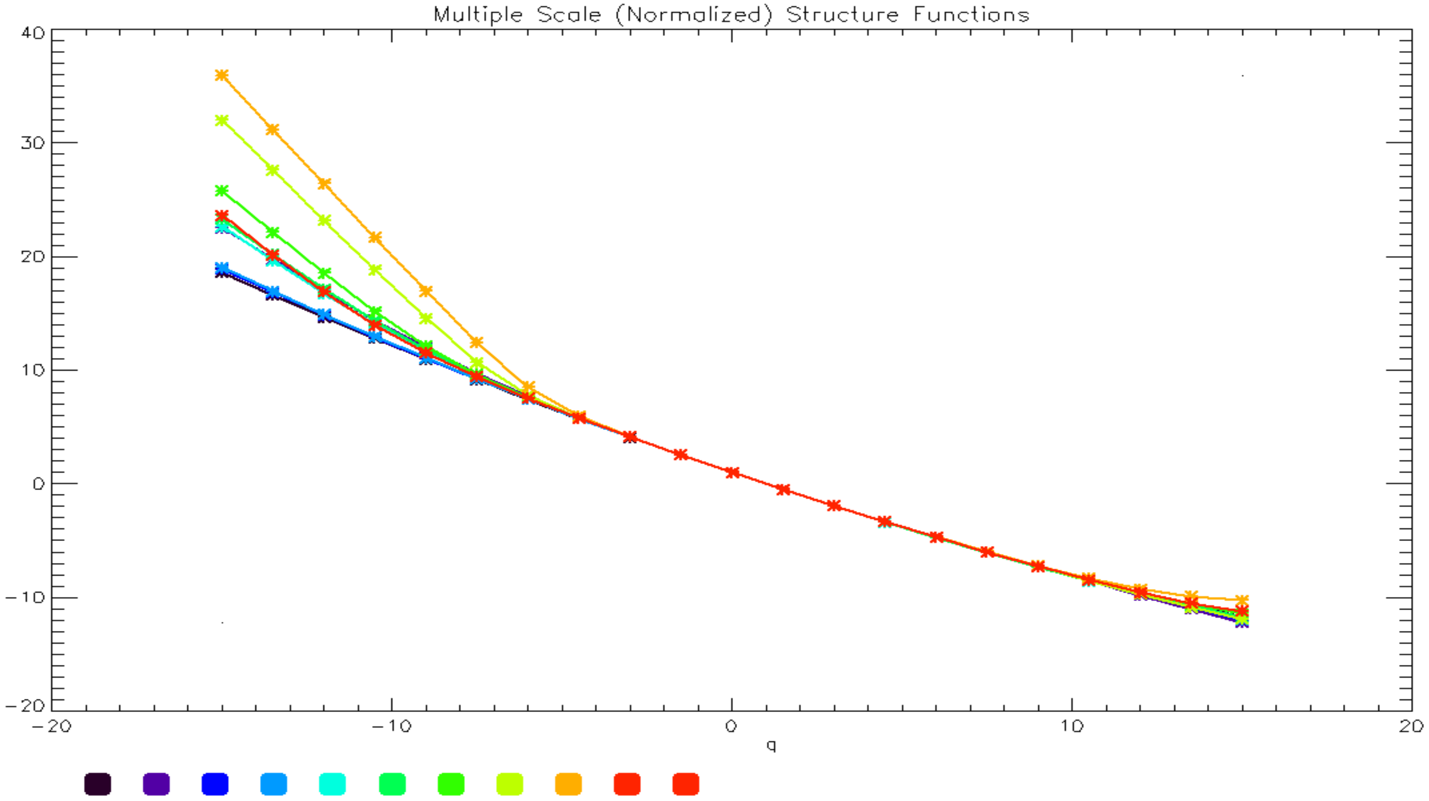
Multiple Scale Structure Functions



Histograms of Variance-Normalized Multipliers



Multiple Scale Structure Functions From Variance-Normalized p's



NAVSEA DAHLGREN Surface Warfare Center Division

Multifractal Spectrum

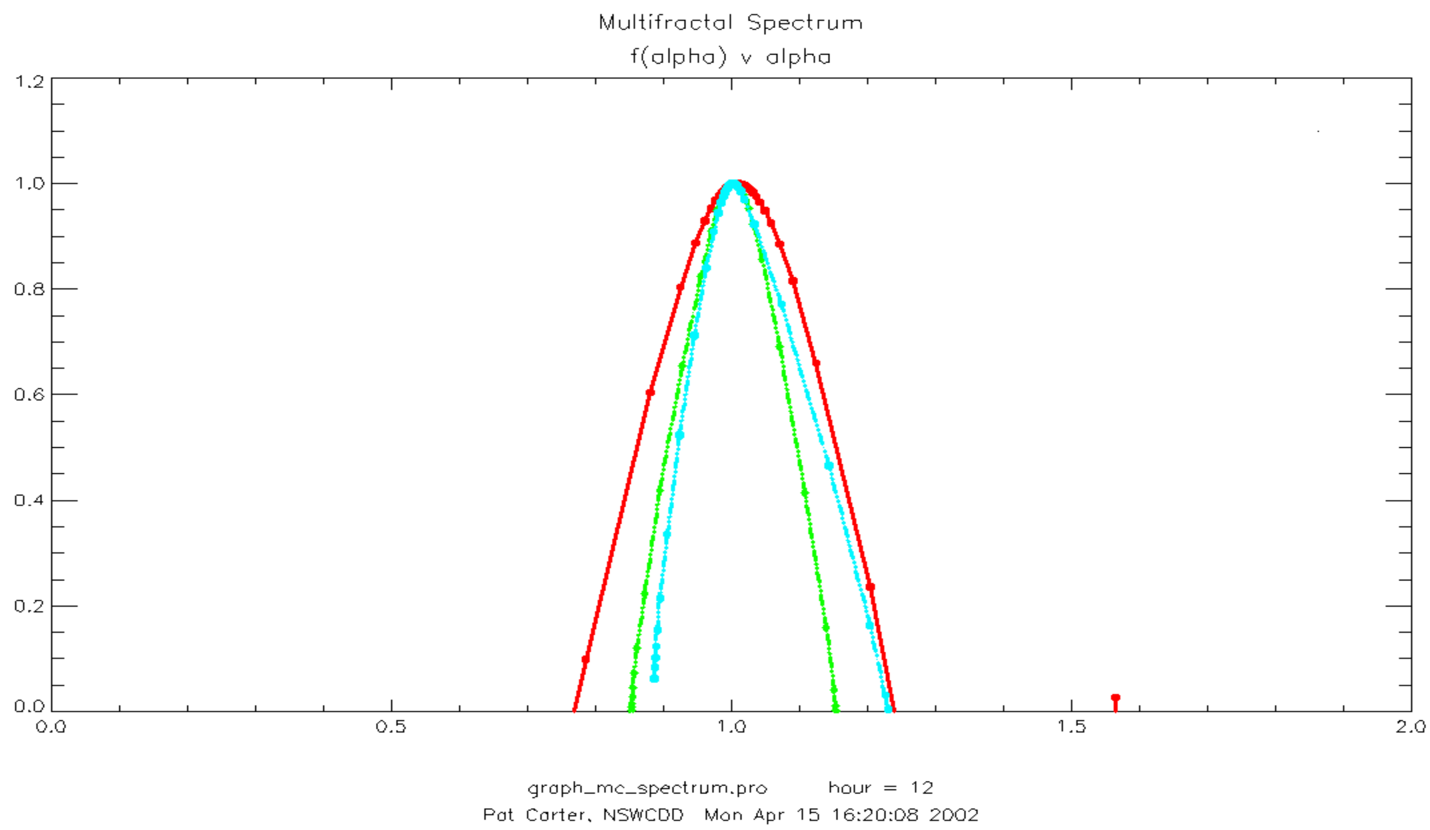
define

$$\tau_\varepsilon(q) = -\frac{\log \sum_i p_i^{\,q}}{\log(\varepsilon)} \qquad \tau(q) = \lim_{\varepsilon \to 0} \tau_\varepsilon(q)$$

"the Legendre transform"

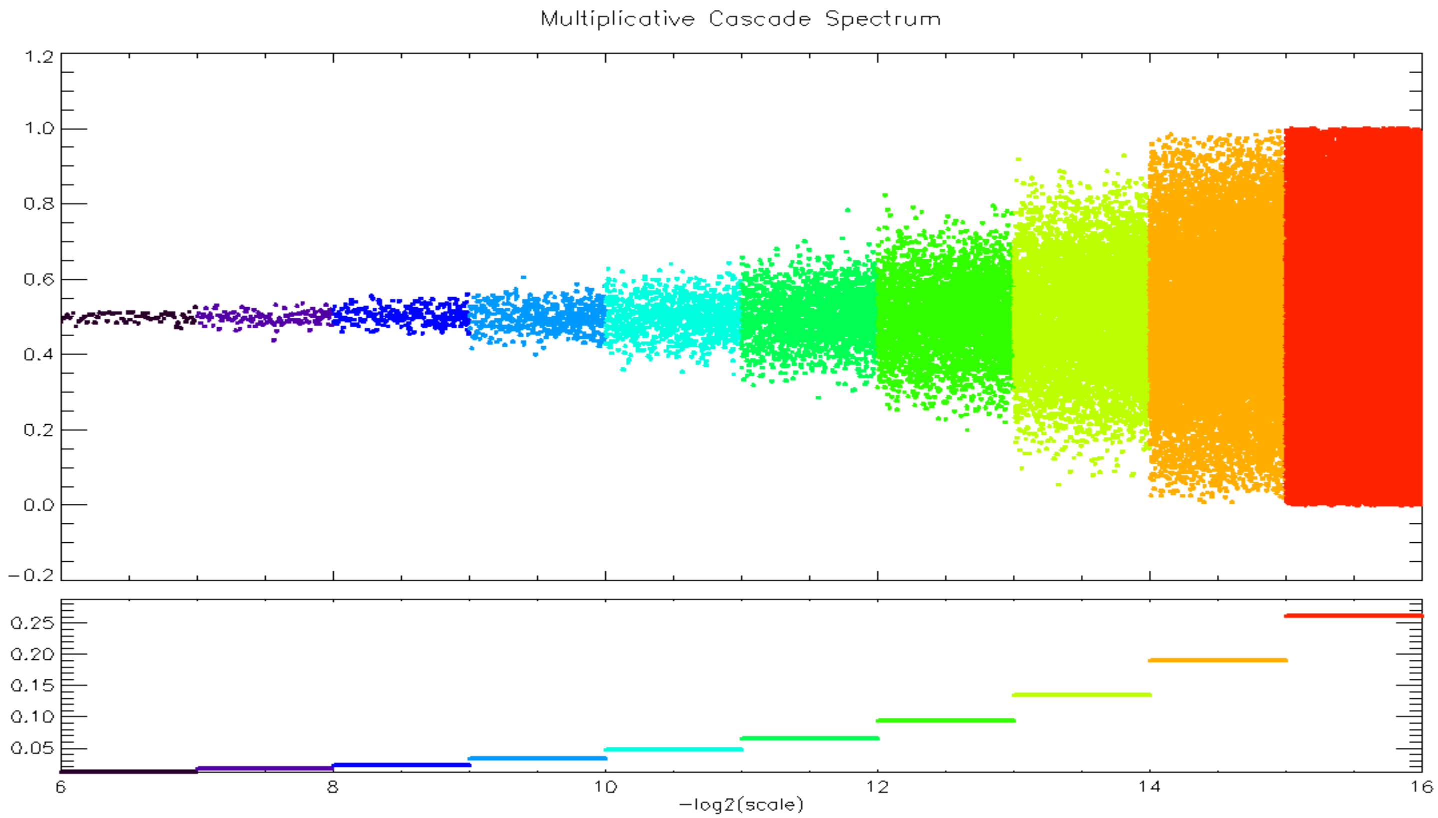
$$\alpha = -\frac{d\tau}{dq} \quad and \quad f = \alpha q + \tau(q)$$

Empirical Approximation to the Multifractal Spectrum - hour 12

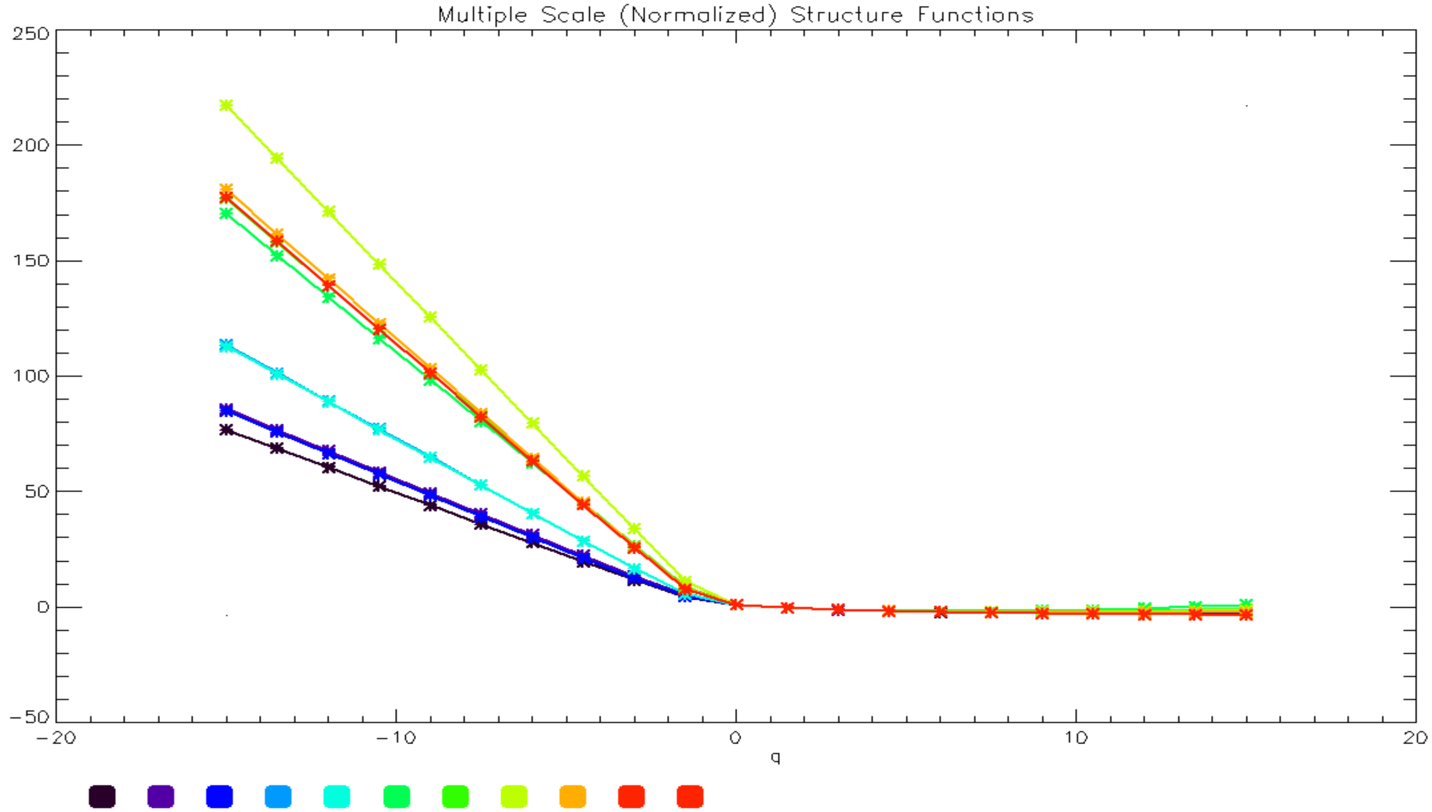


F(alpha) v alpha

Multiplicative Cascade Spectrum
Abs fbm

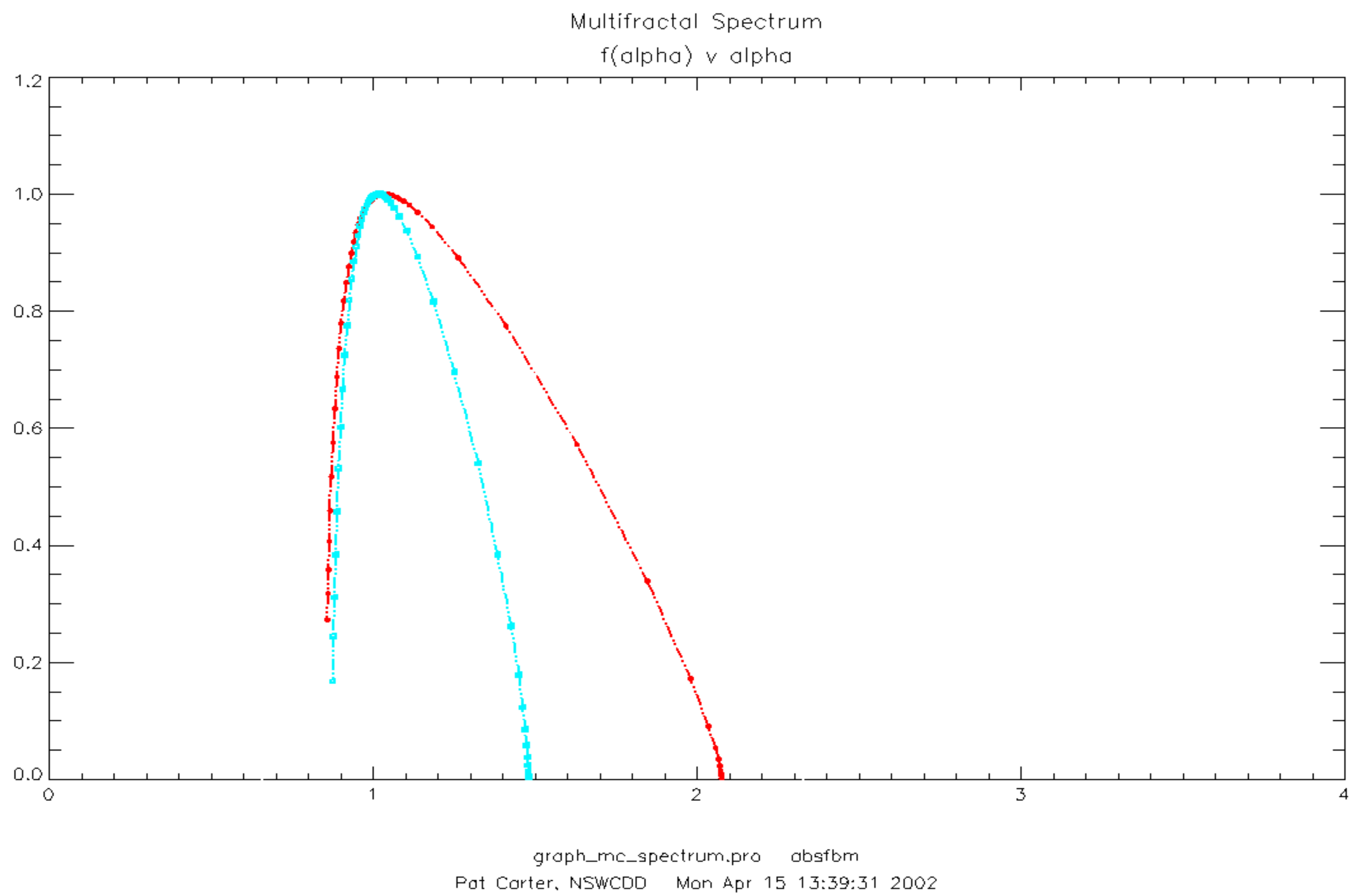


Multiple Scale Structure Functions – abs FBM

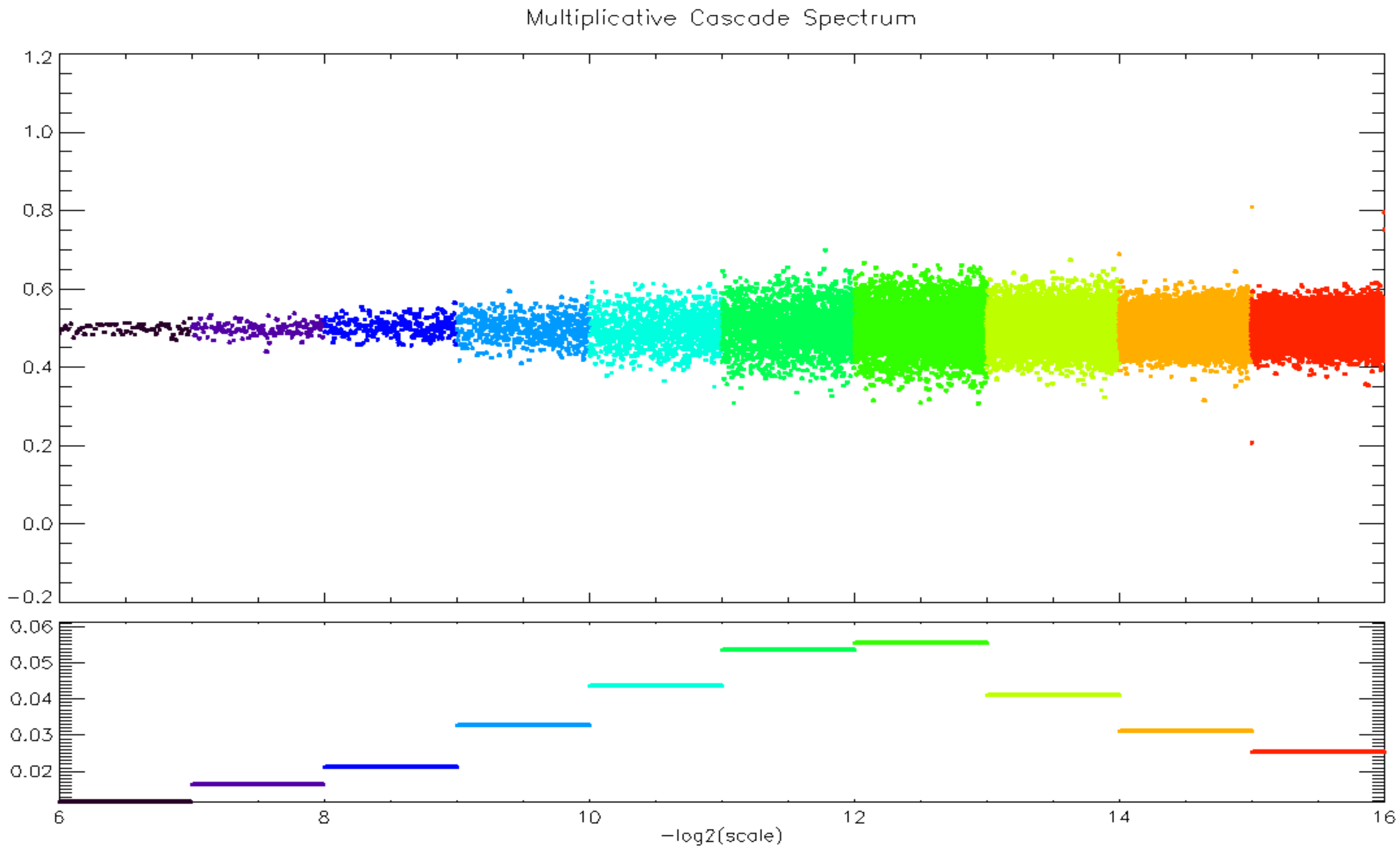


NAVSEA DAHLGREN
Surface Warfare Center Division

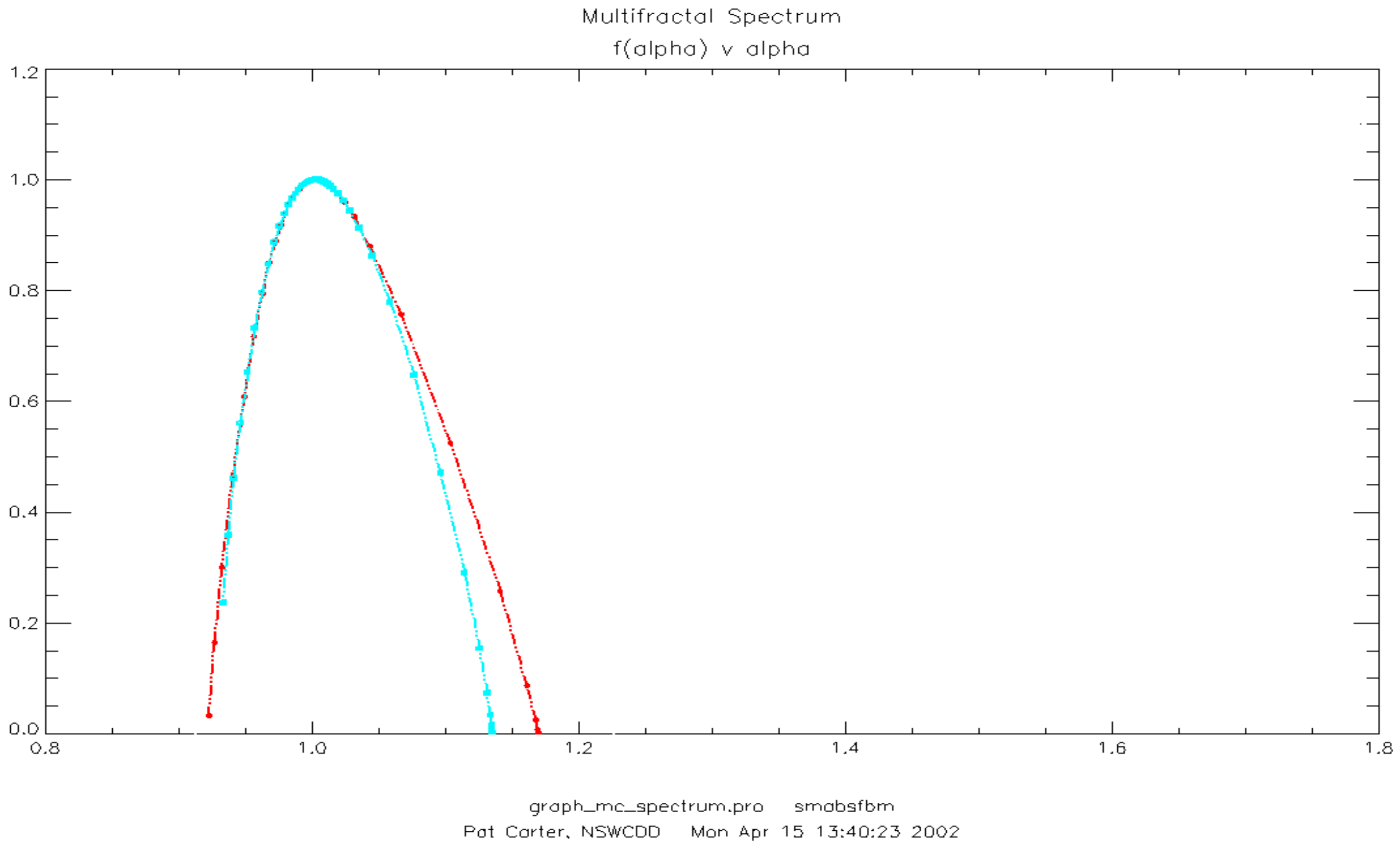
Empirical Approximation to the Multifractal Spectrum – abs fbm



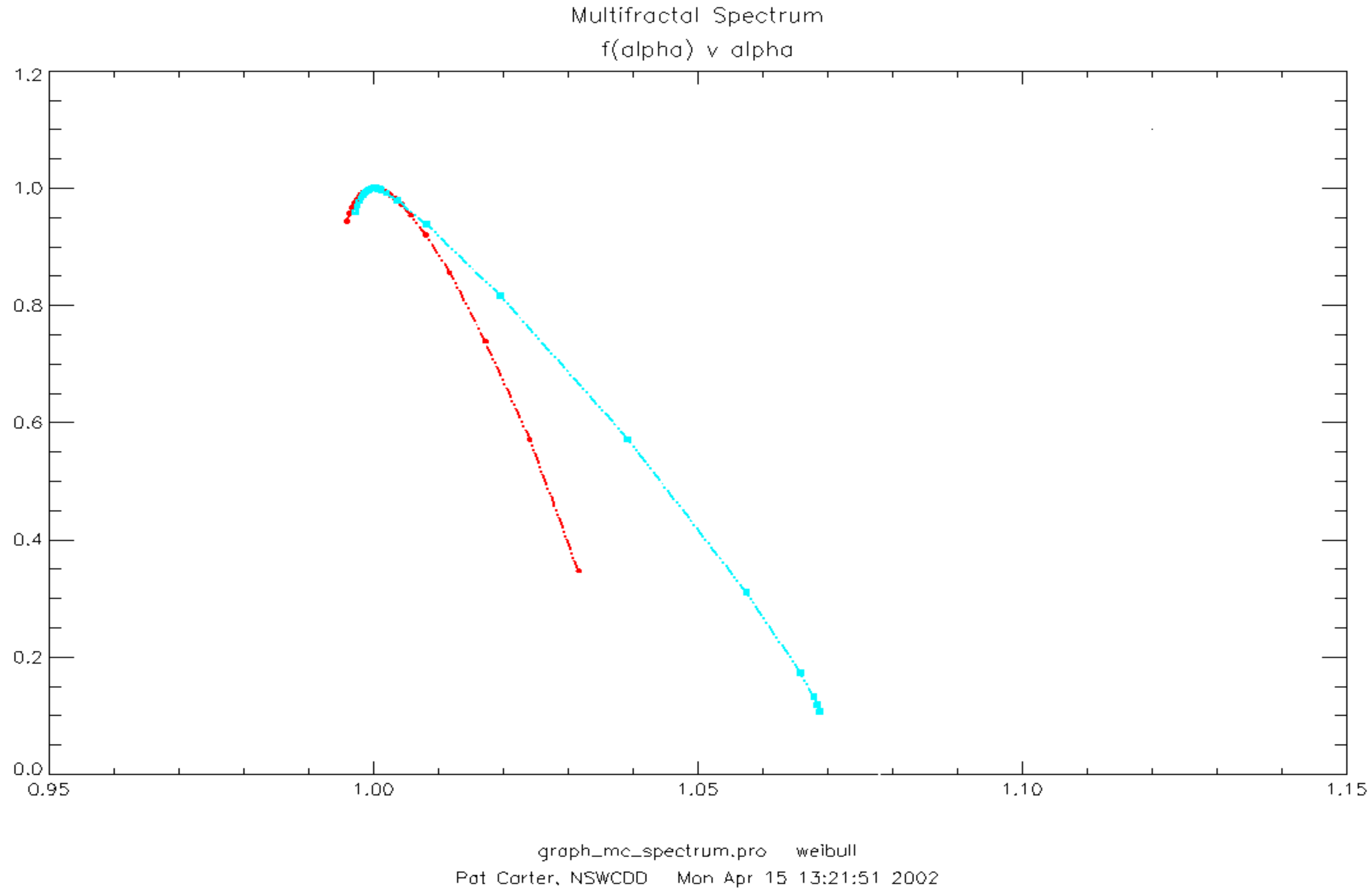
Multiplicative Cascade Spectrum
Abs fBm with smoothing



Multifractal Spectrum – Smoothed fBm



Empirical Approximation to the Multifractal Spectrum – Weibull



NAVSEA
DAHLGREN
Surface Warfare Center Division

Observations/Conclusions I

The packet rate and the multiplicative cascade are a transform pair.

The multiplicative cascade is an appropriate model:

packet rate is positive, so can be interpreted as a measure

over many scales of resolution it is approximately log normal

The multiplicative cascade is an **useful** model if it can be implemented in with a small number of parameters determined by the data of interest:

if the log of the variance is linear in log scale then

the variance is determined by two parameters

at each scale the multipliers can be modeled via a one parameter family, e.g., the symmetric beta distribution - the one parameter is

a function of the variance

Observations/Conclusions II

The multiple scale structure function was useful; the multifractal spectrum was not so useful.

The visualization of the multipliers in time and scale does not convey a lot – the time domain visualization conveys more information