bility and Statistics (E. Eberlein and M. S. Taqqu, eds.) 137–165. Birkhäuser, Boston.
TAQQU, M. S. (1988). Self-similar processes. In *Encyclopedia of Statistical Sciences* 8 352–357. Wiley, New York.
TERRIN, N. and TAQQU, M. S. (1991). Convergence in distribution of sums of bivariate Appell polynomials with long-range dependence. *Probab. Theory Related Fields* **90** 57–81.
VERWAAT, W. (1987). Properties of general self-similar processes. *Bull. Inst. Internat. Statist.* **52** 199–216.
WHITTLE, P. (1956). On the variation of yield variance with plot size. *Biometrika* **43** 337–343.
WHITTLE, P. (1962). Topographic correlation, power-law covariance functions, and diffusion. *Biometrika* **49** 304–314.
YAJIMA, Y. (1985). On estimation of long-memory time series models. *Austral. J. Statist.* **27** 303–320.
YAJIMA, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Ann. Statist.* **16** 791–807.
YAJIMA, Y. (1989). A central limit theorem of Fourier transforms of strongly dependent stationary processes. *J. Time Ser. Anal.* **10** 375–383.
YAJIMA, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Ann. Statist.* **19** 158–177.

Comment: Short-Range Consequences of Long-Range Dependence

Arthur P. Dempster and Jing-Shiang Hwang

We welcome Jan Beran's informative sketch of the history of long-range dependence in many fields of applied statistical science, and likewise his review of the results of several decades of work by mathematical statisticians, mainly on asymptotic sampling theory of various robust as well as normality-based efficient estimators.

Our experience has been with applications of the models, most recently in Hwang (1992) and Dempster and Hwang (1992), to simultaneous estimation of employment time series of 51 U.S. states (including DC) given short input time series of $n = 48$ months. Since our data are fixed, we have emphasized issues related to modeling both time series of sampling error, which a priori have no long-range dependence (ignoring biases that cannot be assessed from our data), and underlying true time series, which appear empirically to have long-range dependence with parameter H close to 1 (but not greater than 1 because nonstationarity of unemployment and employment rate series is a priori implausible).

For inference about the true series, we have emphasized Bayesian thinking, and associated computational issues related to likelihoods of our fixed data, always under assumptions of normality, which appear generally to be reasonable in our case study. Although our theoretical approach to statistical inference is very different from that of Beran, we agree with his opening remarks about dangers from behaving as though traditional ways of thinking about level and variability of underlying short-memory stationary time series models continue to hold in the presence of stationary long-memory models. We direct our brief comments to exposing a few basic small n distinctions between inferences appropriate in situations characterized by short-range dependence and those with long-range dependence. We begin by exhibiting artificially generated pseudorandom "time series" that render in graphical form the main points about estimating the mean and variance of fractional Gaussian noise (fGn) data. We have found it convenient to use alternative notation τ^2 and d in place of σ^2 and H, where $d = 2H - 1$ and τ is chosen so that the spectral density

$$f(\lambda) \sim \tau^2 \lambda^{-d}$$

for λ close to zero. On this scale, $-1 < d < 1$ defines the range for fGn, but $0 < d < 1$ is the range of interest for long-range phenomena, with $d = 0$ corresponding to white noise and $d = 1$ marking the upper boundary where the spectral density first becomes nonintegrable at zero frequency. We use the same frequency domain conventions as Beran, namely, that $-\pi < \lambda < \pi$ and that $f(\lambda)$ is scaled so that σ^2 is its average value with respect to uniform measure. Appropriate roles for the alternative scale parameters τ^2 and σ^2 are elaborated below.

Figure 1 displays four series, each of length $n = 64$, simulated from four different fGn models with $d = 0.8$, 0.9, 0.99, 0.999. Part of the reason for the near coincidence of the curves apart from their levels is that all four were generated from innovations based on the same 64 normal pseudorandom values. In addition, however, the similarity implies covariance structures with remarkably similar forecast operators and resid-

Arthur P. Dempster is Professor and Jing-Shiang Hwang is Postdoctoral Fellow at Department of Statistics, Harvard University, Cambridge, Massachusetts 02138.