JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 92, NO. B1, PAGES 633–648, JANUARY 10, 1987

On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions

ALAN D. CHAVE[1]

Earth and Space Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico

DAVID J. THOMSON

AT&T Bell Laboratories, Murray Hill, New Jersey

MARK E. ANDER

Earth and Space Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico

Robust estimation of power spectra, coherences, and transfer functions is investigated in the context of geophysical data processing. The methods described are frequency-domain extensions of current techniques from the statistical literature and are applicable in cases where section-averaging methods would be used with data that are contaminated by local nonstationarity or isolated outliers. The paper begins with a review of robust estimation theory, emphasizing statistical principles and the maximum likelihood or M-estimators. These are combined with section-averaging spectral techniques to obtain robust estimates of power spectra, coherences, and transfer functions in an automatic, data-adaptive fashion. Because robust methods implicitly identify abnormal data, methods for monitoring the statistical behavior of the estimation process using quantile-quantile plots are also discussed. The results are illustrated using a variety of examples from electromagnetic geophysics.

INTRODUCTION

Reliable estimation of power spectra for single data sequences or of transfer functions and coherences between multiple time series is of central importance in many areas of geophysics and engineering. While the effects of the underlying Gaussian distributional assumptions on such estimates are generally understood, the ability of a small fraction of non-Gaussian noise or localized nonstationarity to affect them is not. These phenomena can destroy conventional estimates, often in a manner that is difficult to detect.

Problems with conventional (i.e., nonrobust) time series procedures arise because they are essentially copies of classical statistical procedures parameterized by frequency. Once Fourier transforms are taken, estimating a spectrum is the same process as computing a variance, and estimating a transfer function is a similar procedure to linear regression. Because these methods are based on the least squares or Gaussian maximum likelihood approaches to statistical inference, their advantages include simplicity and the optimality properties established by the Gauss-Markov theorem [e.g., *Kendall and Stuart*, 1977, chapter 19]. For example, linear regression yields the best linear unbiased estimate when the errors are uncorrelated and share a common variance; this holds independent of any distributional assumptions about them. If, in addition, the residuals are drawn from a multivariate normal probability distribution, then the least squares result is also a maximum likelihood, fully efficient, minimum variance estimate. In practice, the regression model is rarely an accurate description due to departures of the data from the model requirements. Most data contain a small fraction of unusual observations or "outliers" that do not fit the model distribution or share the characteristics of the bulk of the sample. These can often be described by a probability distribution which has a nearly Gaussian shape in the center and tails which are heavier than would be expected for a normal one, or by mixtures of Gaussian distributions with different variances.

Two forms of data outliers are common: point defects and local nonstationarity. Point defects are isolated outliers that exist independent of the structure of the process under study. Typical examples include dropped bits in digital data, transient instrument failures, and spike noise due to natural phenomena (e.g., lightning). Local nonstationarity means a departure from a stationary base state that is of finite duration and must be differentiated from complete nonstationarity, in which the concept of a spectrum must be reformulated [e.g., *Priestley*, 1965; *Martin and Flandrin*, 1985]. A geophysical example of local nonstationarity is seen in observations of the time-varying geomagnetic field: most of the time the data statistics are approximately constant, but this stationary process is interrupted sporadically by brief but intense disturbances such as magnetic storms with markedly different characteristics. In some studies these events are regarded as contaminating noise, and they must be removed to study the underlying process. The influence of these types of outliers on regression problems can be complicated, as aberrant data in the dependent and independent variables produce quite

[1]On leave from Institute of Geophysics and Planetary Physics, University of California, San Diego, La Jolla; now at AT&T Bell Laboratories, Murray Hill, New Jersey.

Paper number 5B5911
0148-0227/87/005B-5911$05.00