Reprinted from

Hirotugu Akaike and Genshiro Kitagawa
Editors

The Practice of Time Series Analysis

Chapter 19

Processing of Time Series Data Obtained by Satellites

Tomoyuki Higuchi
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan
higuchi@ism.ac.jp

19.1   Introduction

Mankind, who are full of intelligent curiosity, are always anxious to satisfy such curiosity and are being engaged in development of new observation instruments. The observation utilizing spacecraft such as rockets, balloons, etc. has aimed at exceeding the limitations of the information that can be measured and obtained exclusively on the ground. The human desire to make probe in a direction of height has been expanded and promoted to the outer space, and furthermore to the universe. Thus commencement of such observation methods as to dispatch observation equipment and apparatus has been seen, that is to say, observation by means of artificial satellites, called satellites simply, has been started.

Satellites that have been dispatched into outer space groping for the farther and farther distance for the purpose of strenuously approaching unknown regions bring about a new viewpoint that has never been experienced by the human beings to the mankind enabling them to make observation of the earth from the outer space, once the eyes of the observation instrument are directed to the earth. When things are viewed in perspective, it is pointed out that there exists a side not only allowing the overall variation to be just glanced at but also permitting various types of the information gathered from individual parts behaving as if they are independent of each other to be united to furthermore deepen the comprehension of the objectives. One of the major objectives of the observation using satellites are also found in such an affair. Satellites have become indispensable instruments as an only measure making it possible to observe a wide range of the earth's surface repetitively and continuously for monitoring destruction of the environment and ecological system of the earth that are now attracting global interest of the people. Satellites' mission at present is greatly expanded from qualitative comprehension of the objectives at the primary stage that is recollected by people in a word of "the probe of the unknown regions"

to quantitative grasp of the objectives taking up and accumulating a large amount of data with prediction as a target.

In extracting the information obtained from satellite data, three types of problems are found. At the first stage, systematic noise with which a possibility to exercise destructive influence on the inference results is pointed out and none of easy separation can be made is included in the data in addition to simple observation noise. Thus the problems to be solved become ill-posed inverse problems. Secondly, it is very difficult to obtain the data of the fixed area on the earth under the same condition as to a local time and weather condition, etc. Therefore when yearly trend is examined, deliberate analysis is required. Thirdly, many of the sensitivity calibration (adjustment of offset, gain, etc.) of the substances and materials that can be equipped in the satellite have to be procured from the surrounding nature for the objectives for comparison owing to the restriction of the weight that can be equipped on the spacecraft, and, strictly speaking, realization of such calibration is virtually impossible. For comparison of the calibration with the data of the different satellites, care should fully be taken to this matter.

From the above description, it is perceived that procedures of analyzing satellite data require the troublesome and ambiguous matters requiring information processing based on human experiments and instinct. Even with the ability of the computers that have been revealing remarkable progress in recent years, it is very difficult to allow human image and concept to be expressed freely and numerically. However researches of the data analysis methods by means of modeling which is designed to make the most use of the information possessed by the data are in progress recently. By adopting a model rich in flexibility that can freely express features of the data, i.e. a model having so many of parameters and by avoiding falling into strict objectivism aloof from reality, the whole frame of the data analysis is constituted based on the fact that the process itself of the human knowledge acquisition contains a kind of subjective factor. For objective identification of the model chosen among so many of possible models, we use an information criterion into which a frame of the Bayesian approach along a line of AIC is taken. Such a data analysis method provides processing of satellite data with new possibility.

For actual analysis, it is necessary that modeling should be tailored to each type of problems. That is to say, the data analysis method by means of a modern Bayesian model becomes made-to-order. In this chapter, we focus on the problems appearing frequently in the time series analysis of the satellite data, and explain the way of how to solve the problems by means of a modern Bayesian model referred to above.

19.2   Problems to be Dealt With

19.2.1   Flow of the Procedures of Satellite Data Analysis

In processing the satellite data, there exist several steps in the flow reaching an intensive operation of the final high-degree/integrated information. However we give simplified description of such processing hereunder. When some data are transferred from a satellite to the ground using an electromagnetic wave, bit errors are liable to be produced due to some causes. First, calibration of the error should be made in the primary processing for decoding operation of received electromagnetic waves (which is called demodulation). Secondly, data conversion in a generalized sense such as transformation of the coordinate system or calibration of the design of hard system of

the observation instrument ahead of launching should be made. Operations to adjust geometric distortion of the image brought about by a variety of causes such as the earth's rotation, curvature, etc., which are called geometric calibration, should also be made (for further details of such operations, refer to Tsuchiya 1990). To the stage of this secondary processing, statistical treatment of the satellite data appears less frequently.

Thirdly, removal of observation noise should be made. "Blot" or "blur" in the images should be removed. The processing that is usually called image processing is the intermediate processing of the above. As observation noise, everything that might possibly prevent the information available at the final stage from being extracted should be included in addition to, needless to say, the noise caused by artificial influence. Vertical, horizontal, or oblique stripes mixed in the images, outliers that were unable to be removed in the primary processing, etc. are exemplified as such a kind of noise. Accordingly the procedure that separates phenomena having evidently nothing to do with the analysis objectives despite the fact of the causes of generation of such phenomena being unknown is also included in this intermediate processing.

The processing at the final stage is high-degree information accumulation corresponding to analysis objects. As processing methods frequently used in terms of the statistical terminology, discriminant analysis, principal component analysis, etc. are designated as representative ones. Detailed interpretation examined from points of view of the individual subjects in the individual scientific fields is applied to the results obtained in the above.

19.2.2 Spin Noise

A regular self-rotational motion (spin) around the spin axis of their body is given to many of satellites for their stabilization of attitude. The spin is likely to bring about cyclical noise to observations and the information is, in some occasions, buried in false information in the worst case. For example when a noise source of which intensity or spatial distribution depends on time is found in a view direction, the observation device measures the noise every time the device's sensor is faced in the direction of the object. As a result, the noise showing an apparent period appears in the obtained data. The noise relating to the spin is just called spin noise for brevity.

Image data from a meteorological satellite to investigate how the cloud is or revealing how natural disaster or human-caused disaster, e.g. forest fires are never free from the influence of this spin noise. Two-dimensional information such as an image is usually composed of combined pieces of one-dimensional information utilizing two-dimensional scans of a normal measurement instrument (Figure 19.1), which is due to the fact that one of scans (in the θ-direction) is usually conducted by the spin motion itself. Even if a scanning has no correlation with the spin motion, the scans themselves induce the noise having periodicity in the data.

None of statistical technique is required with the elimination of the spin noise if the spin noise repeats highly regular cyclic pattern, but it shows a highly temporal/spatial pattern depending on when and where the observation is made. To deal with such a problem, an ad-hoc treatment has been made to allow each data sets to be processed with the instinct based on the application of the existing noise elimination method to so many of the data sets and an abundantly experienced leader's penetrative ability as a judgment criterion. None of sufficient recognition is made at present with not only on-line processing for spin noise elimination but also with a possibility of high-speed large-amount batch processing.
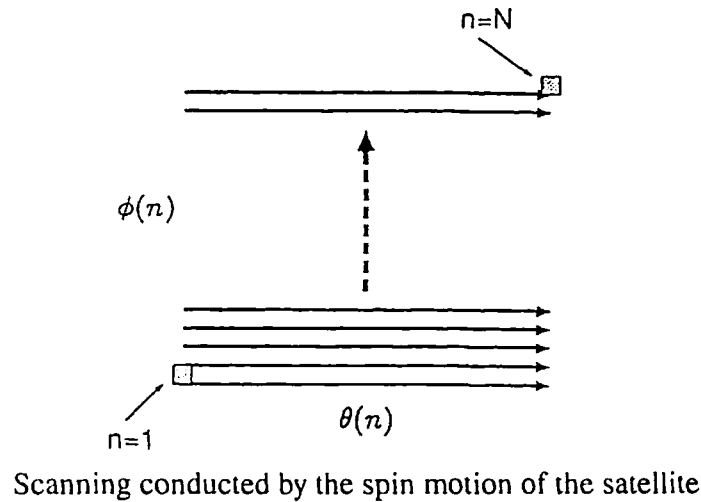
Figure 19.1   A method of the two-dimensional scan to construct an image

In this chapter, elimination of the spin noise classified in the intermediate process-ing is dealt with as a problem accompanied with the analysis of the time series data obtained by an satellite. Descriptions concerning the above is made as a problem of time series as follows. Unless otherwise specified, let the data y_n $(n = 1, \ldots, N)$ be scalar variable. N is the number of the data. Let the angle around the spindle of the self-rotational motion (called a spin axis) be described with θ_n. The angle θ_n is usually assumed to contain none of errors in the measurement values. Even when the angle is not given as data, we assume that the angle can be defined as a function of n with a few parameters. For example, the angle is given as $\theta_n = \theta_0 + 2\pi n \Delta t / T$ with θ_0 as a parameter representing the initial angle. Here, T is the spin period, and Δt is sampling time.

A problem is summarized such that true signals should be estimated after the spin noise s_n as systematic noise is estimated from the data (y_n, θ_n) $(n = 1, \ldots, N)$ comprised of the scalar-variable observation value y_n and θ_n and is removed from y_n.

19.3   Approach by a Bayesian Model –Simple Model–

Here, we deal with a simple model to eliminate spin noise to give an outline of the satellite data analysis by a Bayesian model.

19.3.1   Composition of an Observation Model

The number of the noise components undesirable in estimating the physical quan-tities that is to be obtained is not always restricted as one. In addition to spin noise, observation noise behaving as the white noise sequences, trend component correspond-ing to the movement of the mean value of the data, etc. are often included in the data. An important matter in eliminating these components are not to proceed the processing in succession in a manner as of elimination of the trend components at first, elimination of the spin noise components at the second stage, and elimination of the white noise at the final stage, but to decompose simultaneously the data into the components that are believed to compose the observation. Only by decompos-ing simultaneously the data, which are complimentarily composed of each other, it becomes possible to do rational elimination of the noise with avoidance of one-sided influence by means of the processing of the other components.

Within a framework of the Bayesian model, we explicitly describe the observations by an observation model. In case that the observation y_n is linearly decomposed into a trend component t_n, spin noise component s_n, and observation noise component w_n, let y_n be formulated by means of the observation model as

$$y_n = t_n + s_n + w_n. \tag{19.1}$$

It is intended that observation y_n is expressed by using the parameters such as t_n, s_n together with the observation noise w_n in an explicit manner.

Describing the observation model in advance using a state vector makes it possible to utilize the algorithm of the Kalman filter for parameter estimation, and convenience is offered thanks to good perspective in the calculation method. The observation model is given as

$$y_n = H_n z_n + w_n, \tag{19.2}$$

using a state vector z_n to express the state of the system at time of n. In case of (19.1), from the system model to be referred to later, the state vector is defined by

$$z_n = [t_n, t_{n-1}, s_n, s_{n-1}]^T. \tag{19.3}$$

The mark T denotes transposition. H_n is $1 \times k$ matrix dependent generally on time n, but a constant matrix $H_n = [1, 0, 1, 0]$ in the example in (19.1). With the linear Gaussian Bayesian model, let the observation noise w_n be one-dimensional normal (Gaussian) white noise with the mean 0 and the variance σ^2.

19.3.2 Composition of the System Model

Suppose the spin noise s_n follows the stochastic difference equation

$$s_n - 2C\, s_{n-1} + s_{n-2} = \xi_n, \tag{19.4}$$

where C is given by $C = \cos(2\pi f_c \Delta t)$ with the spin frequency f_c ($f_c = 1/T$). Furthermore, assume that ξ_n is a one-dimensional normal white noise sequences with the mean 0 and variance τ^2. If $\tau^2 = 0$, then the right-hand side of (19.4) is always 0, bringing about a solution of the equation as a sinusoidal wave, i.e. $s_n = A\sin(2\pi n f_c \Delta t + b)$. Here, A indicates an amplitude, whereas b denotes an initial phase. A behavior of s_n seems far from the sinusoidal wave as τ^2 becomes greater. Equation (19.4) becomes a model of the signal so as to be locally a sinusoidal wave with the frequency f_c. Even if the amplitude is changing gradually as time goes or even if the phase is changed in the course, such a signal (hereafter called a cyclic signal for brevity) can satisfactorily be expressed by (19.4) so long as no change is seen only with the frequency. By setting f_c as the spin frequency this model is one adequate enough to eliminate cyclic signals having the performance of the time-dependent amplitude synchronized to the spin frequency.

In case that $C = 1$, that is to say, $f_c = 0$, the left hand side of (19.4) becomes second-order difference of s_n. This is a signal varying smoothly in a time-dependent manner, that is to say, one of the models to represent the trend component. As for (19.1), the model of the trend component t_n is given by

$$t_n - 2\, t_{n-1} + t_{n-2} = \varepsilon_n, \tag{19.5}$$

where ε_n is a one-dimensional normal white noise sequences with the mean 0 and the variance ν^2. In this occasion, the magnitude of ν^2 determines the smoothness of t_n.

As a model of the trend, the one taking the left-hand side as the first-order difference or higher order difference as well can be considered.

In the framework of the Bayesian model, probability distribution is assumed also for the parameters involved in expressing the observation as shown in (19.4) and (19.5). Equations (19.4) and (19.5) give a system model. The parameter appearing in the system model such as τ^2 or ν^2 is called a hyper-parameter.

The system model is expressed as

$$z_n = F_n z_{n-1} + G_n v_n , \tag{19.6}$$

by using a state vector z_n displaying the state of the system. Together with the above,

$$F_n = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 2C & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad G_n = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \tag{19.7}$$

where z_n is given in (19.3). Furthermore $v_n = [\varepsilon_n, \xi_n]^T$ is a Gaussian distribution with mean vector 0 and the variance covariance matrix

$$\begin{bmatrix} \nu^2 & 0 \\ 0 & \tau^2 \end{bmatrix} . \tag{19.8}$$

19.3.3 Kalman Filter

Combination of (19.2) and (19.6) is called a state space representation of the observations. If state space representation is given, then the estimated value $\hat{z}_n$ of the state vector z_n can be obtained by using Kalman filter and smoother algorithm under the fixed values of hyper-parameters. An optimal value of hyper-parameters can be obtained by AIC minimization. In actual application, the initial value $z_{0|0}$ of the state vector and its variance covariance matrix $V_{0|0}$ should be given. Here, $z_{n|j}$ and $V_{n|j}$ express the mean value of the state vector at the time n and variance covariance matrix under the situation where the data $Y_j = [y_1, \ldots, y_j]^T$ are observed. Usually, it is sufficient to think of a diagonal matrix having 0 vector as $z_{0|0}$, a large value (e.g., 10^6 times the variance of y_n) as $V_{0|0}$ for the sake of convenience.

19.3.4 Application to Rocket Data

We show in Figure 19.2 the results* obtained by applying the procedure based on the aforementioned simple model to the data obtained by a rocket which are subjected to the influence of the spin noise as with the case of an satellite. The data provide the observation $I(z)$ of the intensity of the 5577Å airglow emitted by the excited oxygen atoms at the altitude (height) z. The quasi-cyclic behavior component seen in Panel (a) indicates the rocket's spin noise. The actual observation $I(z)$, is not, the intensity $J(z)$ of the light emitted by the oxygen atoms in the vicinity of the height (altitude) z, but the integrated quantity $I(z) = \int_z^{+\infty} J(z')dz'$ ranging from the infinity to the height (altitude) z. To obtain the estimate of $J(z)$ from $I(z)$, it is necessary, first of all, to eliminate the spin noise together with the observation noise from $I(z)$.

In Figure 19.2(b), the three components t_n, s_n and w_n obtained by decomposing $I(z)$ are shown. The components s_n and w_n are the terms corresponding respectively
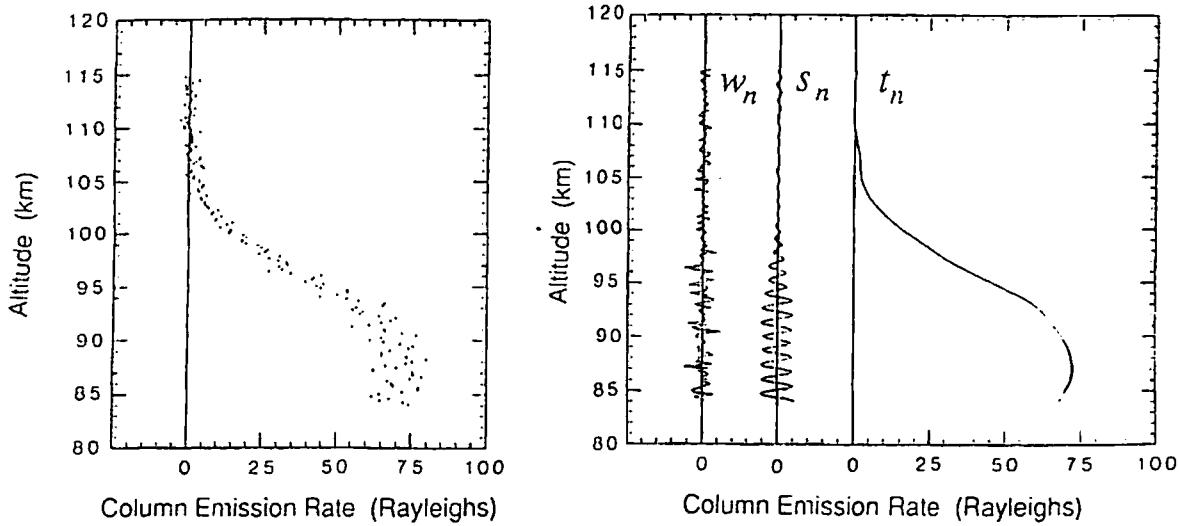
Figure 19.2 (a) Height profile of the intensity of the 5577Å atmospheric light emitted by the excited oxygen atoms. The ordinate is taken as the altitude direction. (b) Profile decomposed into the individual components of the trend t_n, spin noise s_n, and observation noise w_n.

to the spin noise and observation noise, and it is possible to estimate the discretized $J(z)$ by taking the first-order difference of the component t_n obtained as a result of elimination of an effect associated with a spin from the observation $I(z)$. It is clear that even such a simple model makes it possible to realize information processing with extremely high practicality. For details, refer to Higuchi *et al.* (1988).

19.3.5 Interpretation as a Linear Filter

An estimation of the trend components by using the Bayesian model is regarded as an equivalent to that of low-frequency components. Let the extraction method of the spin noise be compared to the trend estimation method by a classical low-pass filter. When the observation model is given by a simple sum of the several components as seen in (19.1), a simple relationship can be given between the estimates of the obtained components and the data. In case of the above example, it is expressed by

$$T_N = L_t^{\alpha^2} Y_N \qquad (19.9)$$

$$S_N = L_s^{\beta^2} Y_N, \qquad (19.10)$$

where $T_N = [t_1, \ldots, t_N]^T$, $S_N = [s_1, \ldots, s_N]^T$, $Y_N = [y_1, \ldots, y_N]^T$. Furthermore $L_t^{\alpha^2}$ and $L_s^{\beta^2}$ are $N \times N$ matrices corresponding to each t_n and s_n determined by the system model and the value of the hyper-parameter. Here, $\alpha^2 = \nu^2/\sigma^2$ and $\beta^2 = \tau^2/\sigma^2$.

To obtain the expression of (19.9) and (19.10) in a frequency domain, executing discrete Fourier transformation gives

$$\widetilde{T_N} = \widetilde{L_t^{\alpha^2}} \, \widetilde{Y_N} \qquad (19.11)$$

$$\widetilde{S_N} = \widetilde{L_s^{\beta^2}} \, \widetilde{Y_N}. \qquad (19.12)$$

The diagonal components of $\widetilde{L_t^{\alpha^2}}$ and $\widetilde{L_s^{\beta^2}}$ determine the frequency response produced when the transformation from the data to each components is viewed as a linear filter.
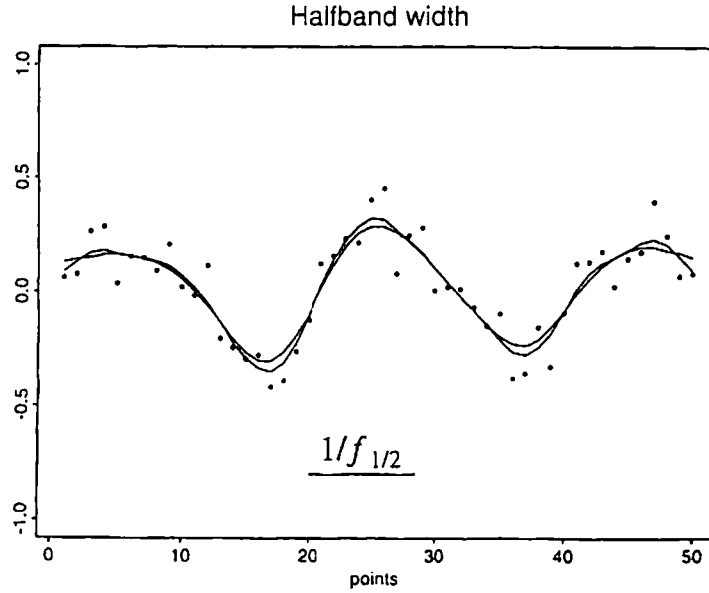
Figure 19.3 Comparison of the smoothing by means of the Bayesian model with the smoothing by means of the moving average. In case of the Bayesian model, the smoothing parameter is automatically determined. $1/f_{1/2}$ is the half-band width obtained from the determined value in accordance with (19.13).

The diagonal component of $\widetilde{L_t^{\alpha^2}}$ describes to what extent of the frequency the component will be allowed to be preferentially passed in the frequency domain, whereas the diagonal component of $\widetilde{L_t^{\beta^2}}$ describes to what extent of the frequency component around f_c is allowed to be selectively passed. The hyper-parameter controls this frequency response.

When the procedure to extract the trend component is regarded as applying a low-pass filter that takes out the low-frequency components, the frequency of which the magnitude of the frequence response attenuates to a half value of that at the frequency 0, i.e. the half-band width $f_{1/2}$ is given approximately as a function of the hyper-parameter α^2 in accordance with the low-pass filter based on the model given by (19.5) (Higuchi 1991)

$$f_{1/2} \simeq \frac{\sqrt{\alpha}}{6} \quad (\text{for } 2^{-12} \leq \alpha^2 \leq 2^2),\qquad\qquad (19.13)$$

where the frequency domain is restricted in a range from 0 to $1/2$ on the assumption that the measurement time interval $\Delta t = 1$. Shown in Figure 19.3 are the artificially-generated data y_n, the trend component estimated based on the hyper-parameter automatically determined in accordance with the AIC minimization procedure, and the result of smoothing by means of the simple moving average

$$\hat{t}_n = \frac{1}{2K^* + 1} \sum_{i=-K^*}^{K^*} y_{n+i},$$

having a half-band width which is given by the frequency response from the above expression. K^* is defined by $K^\dagger$ where the frequency response function $H_K(f_{1/2})$ of $(2K + 1)$-point moving average

---

†It is possible to give in a manner as of $K^* = [1/(4f_{1/2}) - 1/2]$, where $[\cdot]$ represents the integer part.

$$H_K(f_{1/2}) = \frac{1}{2K+1}\left[1 + 2\sum_{n=1}^{K}\cos(2\pi n f_{1/2})\right] \tag{19.14}$$

takes the nearest value to 0.5. The value of the optimal hyper-parameter obtained for the data shown in Figure 19.3 is $\alpha^2 = 0.25$, and it turns out that $K^* = 2$. From the fact that the trend component almost shows a good agreement with the moving average, $f_{1/2}$ given by (19.13) is appropriate enough.

19.4 Example of a Simple Model

In this paragraph, we show several models having a possibility for application to satellite data processing in a concrete manner to exemplify how to construct models.

19.4.1 Extension to Multi-Components

When a signal is composed of several cyclic components, two models can be considered. One is the model to generalize the observation model as

$$y_n = s_n^1 + s_n^2 + \cdots + s_n^M + w_n, \tag{19.15}$$

where s_n^m is the m-th quasi-cyclic (wavy) component, and as in the case of (19.4), the system model for each component is given as

$$s_n^m - 2C^m s_{n-1}^m + s_{n-2}^m = \xi_n^m. \tag{19.16}$$

Furthermore, C^m is a constant given by $C^m = \cos(2\pi f_c^m \Delta t)$ in case that the frequency of s_n^m is f_c^m. The variance τ_m^2 $(m = 1,\ldots,M)$ of ξ_n^m is also a hyper-parameter for this model.

With another one, the system model is given as

$$\sum_{j=0}^{2M} a_j s_{n-j} = \xi_n, \tag{19.17}$$

where the coefficient a_j is obtained by comparing the coefficients of s_{n-j} in the equality

$$\sum_{j=0}^{2M} a_j s_{n-j} = \prod_{m=1}^{M}\left(1 - 2C^m B + B^2\right)s_n, \tag{19.18}$$

where B is the backward operator defined by $s_{n-1} = Bs_n$.

19.4.2 Model of the Decaying/growing Wave

Let it be considered that s_n is a wavy signal locally decaying or growing. When it can be supposed that $s_n = A\exp\{2\pi(g_c + if_c)n\Delta t + ib\}$, the model in (19.4) is generalized as

$$s_n - 2\gamma_c\cos(2\pi f_c)s_{n-1} + \gamma_c^2 s_{n-2} = \xi_n, \tag{19.19}$$

where $\gamma_c = \exp(2\pi g_c)$, and the positive g_c is called a growth rate of the wave and the negative one is called a damping rate. When $g_c = 0$, (19.19) evidently coincides with (19.4). The model in question expresses the free oscillation in case that friction is found, and is adequate enough to express a wavy behavior globally growing or decaying. For an extension to multiple components, it is recommended that (19.19) be used instead of (19.4) in the discussion in the previous subsection.

19.4.3   Seasonal-Adjustment-Type Model

Long-term economic activity should be investigated in the time series analysis of the economic data, and therefore seasonal adjustment to remove the variation concerned with the season from the data is a very important subject for researches. Because the seasonal adjustment is to remove the variation having one-year cycle from the data, the Bayesian model proposed for the seasonal adjustment can be directly applied to spin noise elimination provided that the cycle of the spin can be given instead of the yearly cycle. When the cycle is r, the simplest model is given as

$$s_n - s_{n-r} = \xi_n . \tag{19.20}$$

In case of monthly data (i.e. the data obtained every month), $r = 12$. This model, which allow s_n to have a non-zero mean component such as $s_n \equiv$ constant, is not an adequate one when the trend component in addition to the spin noise component is intended to be extracted from the data. When (19.20) is re-written as

$$\begin{aligned}
s_n - s_{n-r} &= (1 - B^r)s_n \\
&= (1 - B)(1 + B + B^2 + \cdots + B^{r-1})s_n \\
&= \xi_n ,
\end{aligned} \tag{19.21}$$

it can be seen that the first part of the second line coincides with the first order difference operator used for the model of trend. It is therefore recommended that the condition required by the second pair of the parentheses should be made the model for seasonal adjustment, i.e.

$$\sum_{j=0}^{r-1} s_{n-j} = \xi_n . \tag{19.22}$$

The seasonal adjustment model can be applied to elimination of the spin noise by putting $r = T/\Delta t$, provided that the spin period T is of multiple length of the sampling interval Δt. With this model, it is utterly unnecessary for the spin noise to be of sinusoidal wave, and any kind of signals can be expressed as a fundamental principle provided that the cycle is r. Needless to say, it is possible for the repeating patterns to be varied slowly dependent on time.

19.5   Point Noise Source Model

We hereunder show the approach where the physical structure of the noise source is explicitly taken into the observation model.

19.5.1   Uni-Variate

Suppose the objective generating noise is of the point noise source. At that time, the noise can be written by the product $f(\cdot)I_n$, where I_n is a time-dependent intensity of the point noise source and $f(\cdot)$ can be assumed to be described as a function of the relative angle θ formed between the point noise source and the instrument. Now assume that the relative positioning between the satellite and the point noise source remains unchanged during the period of the observation, then we can get $f(\cdot) = f(\theta)$.

Under the assumption of such a noise source, let a problem to estimate the physical quantity $x = x(t, \theta)$ depending on the time t and the space (which is equivalent to the angle θ in this occasion) be considered. Now suppose that the time dependency
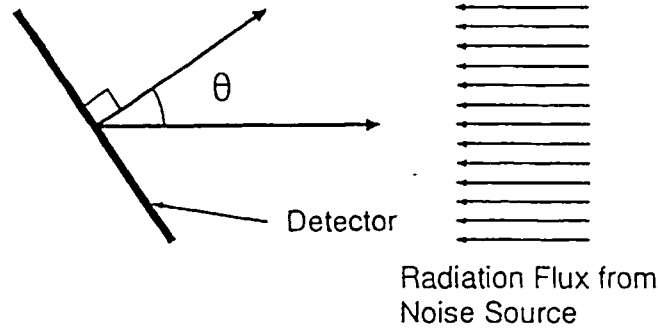
Figure 19.4 Exemplified schematic relation between the observation and the instrument. The observation is received by the sensor of parallelly-placed plates. Around the satellite, the observation is in proportion to the flux from light sources almost parallel with each other in the distance.

of x is almost negligible within Δt ($|\Delta t \cdot \partial x / \partial t| \simeq 0$). Also suppose that the x within this interval is a physical quantity determined exclusively by the angle. That is to say, suppose that the component x_n obtained by removing an effect of the point noise source and observation noise from the data y_n at time $t = n$ is given by $x_n \simeq x(\theta_n)$ using the angle $\theta = \theta_n$. Furthermore, by using a small angle $\Delta\theta$, let $x(\theta + \Delta\theta)$ be approximated as

$$x(\theta + \Delta\theta) \quad \simeq \quad x(\theta) + x(\dot\theta) - x(\dot\theta - \Delta\theta). \tag{19.23}$$

At that time, assuming that the θ_n is so designed as to allow $\Delta\theta = \theta_n - \theta_{n-1}$, we have

$$x(\theta_{n+1}) \quad \simeq \quad x(\theta_n) + x(\theta_n) - x(\theta_{n-1})$$
$$x_{n+1} \quad \simeq \quad 2x_n - x_{n-1} . \tag{19.24}$$

In short, a trend model is adopted for x_n assuming that the second-order difference shows small variation.

The intensity I_n of the point noise varies slowly dependently on time, and we adopt a trend model assuming that the first-order difference shows small variation. At that time, $z_n = [I_n, x_n, x_{n-1}]^T$ with respect to the state vector, whereas H_n of the observation model is given as $H_n = [f(\theta_n), 1, 0]$. As $f(\theta)$, let a function having small parameters be considered corresponding to the purpose of the application. For example, when the obtained data are so designed as to be in proportion to the flux coming out from the noise source, we obtain (Figure 19.4)

$$f(\theta) = \begin{cases} \cos\theta & |\theta| < 90° \\ 0 & 90° \le |\theta| \le 180°. \end{cases} \tag{19.25}$$

In addition to the above, the factors such as $\cos^2\theta_n$, $\exp(-\sin^2\theta_n)$, quadratic polynomial, etc. are considered as candidates. $f(\theta)$ should be designed so as to express the characteristics of the instruments in dealing with practical problems.

19.5.2 Expansion to Multi-Variate

To examine space-distribution structure of the object, several instruments steered into different directions are occasionally mounted on a plane onto which a spin axis is included. At that time, the observation becomes the multi-variate $y_n = [y_n^1, y_n^2, \ldots, y_n^M]^T$.

Furthermore, y_n^m is the observation at time n obtained with the m-th instrument (channel-m). Here, M is called the channel number. When the angle of the channel m from the spin axis in the plane including the spin axis is described as ϕ^m, the relationship $f(\cdot)$ between the channel m and the point noise source is written by θ_n and ϕ^m. That is to say, $f(\cdot) = f(\theta_n, \phi^m)$. At that time, not by considering a model with respect to every channels, that is, for y_n^m of the individuals but by thinking of an extensive model for multi-variate time series y_n, an estimation of a function form of $f(\theta_n, \phi^m)$ can be made based on the more rational assumption. Furthermore, it becomes possible to estimate features such as the difference among the equipment characteristics among channels from the data.

Suppose the component x_n^m which is obtained by removing the influence of the point noise source at time n and the observation noise from the observation of the channel m along line of the solving method in case of uni-variate, is in accordance with the second-order difference model. On the other hand, suppose that I_n is expressed by the first-order difference model as in the case of uni-variate. From this it is understood that the state vector becomes the $(1+2M) \times 1$ vector $z_n = [I_n, x_n^1, x_{n-1}^1, \ldots, x_n^M, x_{n-1}^M]^T$. On the other hand, H_n of the observation model becomes the $M \times (1+2M)$ matrix

$$\begin{bmatrix} f(\theta_n, \phi^1) & 1 & 0 & & & & 0 \\ f(\theta_n, \phi^2) & & 1 & 0 & & & \\ \vdots & & & & \ddots & & \\ f(\theta_n, \phi^M) & 0 & & & & 1 & 0 \end{bmatrix}. \tag{19.26}$$

The structure of the variance/covariance matrix of the vector of the observation noise $w_n = [w_n^1, w_n^2, \ldots, w_n^M]^T$, should be determined by taking into account the factors including the characteristics among the channels.

19.5.3 Expansion of the Point Noise Source Model

When the influence from the noise source cannot be expressed by the point noise source referred to above, it is necessary to express $I_n \cdot f(\theta_n)$ in a form with higher flexibility. Let it be considered that the θ dependency of the spin noise should be expressed by a first-order spline function $R_n(\theta)$ at each time n. First of all, let node points θ_j ($j = 1, \ldots, J$) be determined by equally dividing $[0, 360°]$ by d_θ. Here, $J = 360/d_\theta$. Secondly, let values of $R_n(\theta)$ at each node points be denoted as r_n^j. $R_n(\theta)$ in a range of each regions $[(j-1)d_\theta, jd_\theta]$ is given by

$$R_n(\theta) = (1-a)r_n^j + ar_n^{j+1}, \tag{19.27}$$

where $a = \theta/d_\theta - [\theta/d_\theta]$ with $[\theta/d_\theta]$ indicating the integer part of θ/d_θ. Also on the assumption that $R_n(\theta)$ is a periodic function, $r_n^{J+1} = r_n^1$.

By assuming that $R_n(\theta)$ varies slowly dependently on time, let it be assumed that r_n^j is in accordance with the first-order difference model. It is possible for the variance of the first-order difference to be in common with j or to be optimized so that the likelihood will individually be maximum. In general, some kinds of prior information are available, and modeling so designed as to effectively utilize such information will be made. As in the case of uni-variate point noise source, assume that x_n to be measured can be expressed by the second-order difference model. Accordingly the state vector becomes $z_n = [r_n^1, r_n^2, \ldots, r_n^J, x_n, x_{n-1}]^T$. H_n of the observation model
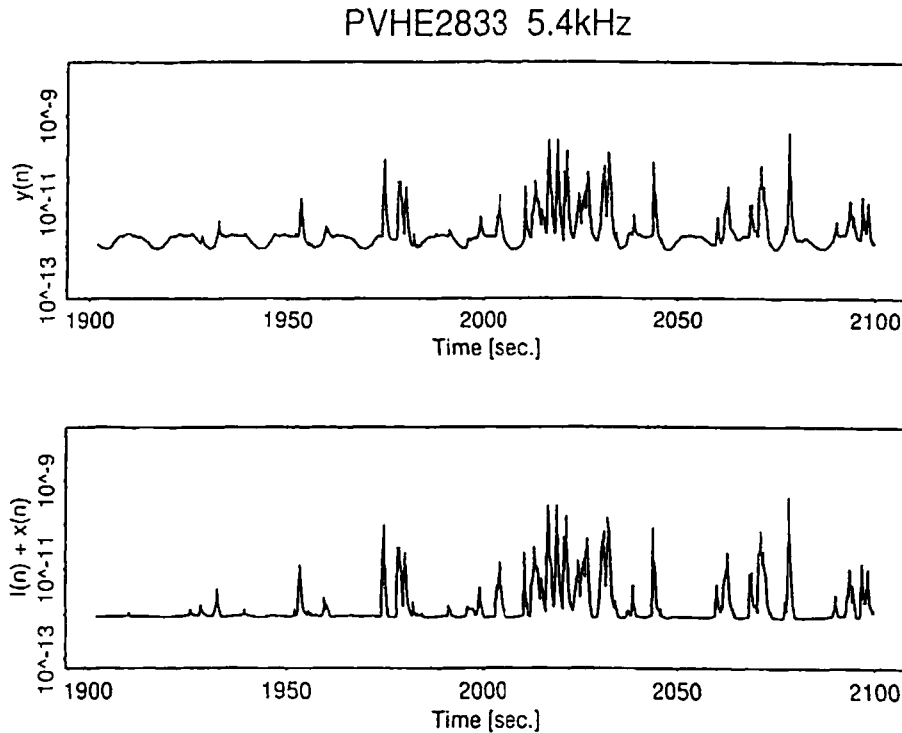
PVHE2833 5.4kHz

Figure 19.5 (a) Intensity of the electric field observed by Pioneer Venus Orbiter. (b) Data after systematic noise caused by spin noise is eliminated.

becomes $1 \times (J + 2)$ vector defined by

$$H_n = [\overbrace{0,\dots,0}^{j_n \text{ times}},1-a,a, \overbrace{0,\dots,0}^{J-j_n-2 \text{ times}} ,1,0], \qquad (19.28)$$

where j_n is an integer defined by $[\theta_n/d_\theta]$.

When there locally exist the generation sources of the signal x to be measured in space, no approximation of (19.23) can be established. In this occasion, we decompose the observation y_n as

$$y_n = w_n I_n f(\theta_n) + x_n , \qquad (19.29)$$

in accordance with selection (by combining the judgment of $x_n = 0$ or no, with estimation of I_n and $f(\theta)$ (Higuchi 1993)).

Assume I_n varies slowly, and is expressed by the first-order difference model. At the same time, assume $f(\theta)$ shows a spatially smooth behavior. These assumptions can be accomplished by applying the trend model taking the second-order difference of f_j ($j = 1,\dots,360/d_\theta$), as an infinitesimal amount, where f_j is defined by discretizing $f(\theta)$ by d_θ. The model in question, which is a model of the observation depending on both time and space, is, in short, one of the simplest statistical model called generally space-time model. Let a search of the optimal x_n, I_n, and f_j be made by the procedure to do independently the judgment as to whether $x_n = 0$ or not, smoothing the time domain to obtain I_n, and smoothing the space domain to obtain f_j.

The results[*] obtained by applying this method to real satellite data are shown in Figure 19.5. Shown in (a) is just part (792 points) of the 20394 data y_n observed.

The cyclic variation is equivalent to the spin noise. When the signal corresponding to x_n is observed, y_n takes extremely large values. In (b), the data $I_n + x_n$ obtained by removing and calibrating the influence of $f(\theta_n)$ from the observed data y_n are shown. It is found that the systematic noise evidently concerned with the spin noise seen in (a) is removed in (b). I_n in the interval shown in this figure is almost constant, but varies when viewed through all of the data.

19.6   Conclusions

In the above, we explain several models to remove spin noise in the satellite time series data. Although the frame of the Bayesian model is extremely of the unified style, it is necessary to think of a model specialized into the individual problems in case of application. To develop a new model for each actual problem takes a time for the users, but the new analysis method based on the new model will enable the analyzers' ideas or knowledge to be effectively utilized to the utmost.

References

Higuchi, T., K. Kita, and T. Ogawa (1988), "Bayesian statistical inference to remove periodic noises in the optical observation aboard a spacecraft," *Applied Optics*, Vol. 27, No. 21, 4514-4519.

Higuchi, T. (1991), "Frequency domain characteristics of linear operator to decompose a time series into the multi-components," *Annals of the Institute of Statistical Mathematics*, Vol. 43, No. 3, 469-492.

Higuchi, T. (1993), "A method to separate the spin synchronized signals using a Bayesian approach," *Proceedings of the Institute of Statistical Mathematics*, Vol. 41, No. 2, 115-130.

Higuchi, T. (1994), "Separation of spin synchronized signals using a Bayesian approach," *Proceeding of The Frontiers of Statistical Modeling: An informational approach*, (eds. H. Bozdogan), Kluwer Academic Publishers, 193-215.

Tsuchiya, K. (1990), *An Introduction to Remote Sensing*, Asakura Publishing Co., Tokyo (in Japanese).