Planet. Space Sci., Vol. 37, No. 11, pp. 1327–1331, 1989
Printed in Great Britain.

0032–0633/89 $3.00+0.00
Pergamon Press plc

BAYESIAN STATISTICAL INFERENCE OF AIRGLOW PROFILES FROM ROCKET OBSERVATIONAL DATA: COMPARISON WITH CONVENTIONAL METHODS

K. KITA, T. HIGUCHI* and T. OGAWA
Geophysics Research Laboratory, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

(*Received* 4 *April* 1989)

Abstract—It is crucial for the airglow investigation to infer an altitude profile of the volume emission rate from the column emission rates observed by using a spacecraft. A new inference procedure based on the Bayesian statistics was proposed and it was applied to the analyses of the rocket airglow data. In this paper, its high performance is demonstrated in comparison with some conventional methods. The procedure has two major advantages: one is the use of statistical criteria to infer the most likely model of the volume emission rates, and the other is high efficiency to eliminate a periodic modulation.

1. INTRODUCTION

Rocket observation has been carried out for measurement of the altitude distribution of the airglow. The physical quantity measured in a rocket photometry observation is the column emission rate which is given by integrating volume emission rates along the line-of-sight. Accordingly, we are concerned with reconstructing an altitude profile of volume emission rate from the observed column emission rates. Column emission rates can be converted to vertical column emission rates in terms of an appropriate geometrical factor as far as the instrumental field-of-view is narrow enough. Since the vertical column emission rate I is related to the volume emission rate J by

$$I(z) = \int_z^{+\infty} J(z')\,dz',$$

where z denotes altitude, we have only to differentiate the vertical column emission rate with respect to altitude in order to obtain the volume emission rate. However, a random fluctuation always contaminates the observed data, and a spurious modulation sometimes occurs in conjunction with a periodic change in the rocket attitude, even if we eliminate the modulations due to extraterrestrial background emissions and geometrical change in look direction to the airglow layer. Since the differential operation is sensitive to these noises, the direct inversion of the observed data leads to a large error. Therefore, a procedure such as curve fitting or smoothing is essential for getting a result consistent with our *a priori* knowledge: a smooth altitude variation. At the same time, it is important to preserve small-scale structures that are aeronomically significant.

For the above purpose, many techniques have been proposed and applied to the analysis of the rocket airglow data. However, in the case that a spurious modulation occurs in the observed data, neither curve fitting nor smoothing techniques can eliminate the modulation without distorting features of the airglow emission rate itself. In addition, in these conventional techniques, we have to set a smoothing parameter arbitrarily in the reconstruction of the volume emission rate profile, and this arbitrariness leads to an uncertainty, which is a serious problem when we are concerned with small-scale structures in the airglow layer.

We have proposed a new method of the Bayesian statistical inference to remove both a random noise and a periodic modulation from the observed data, and demonstrated its availability by applying it to synthetic data sets (Higuchi *et al.*, 1988). An information criterion on goodness-of-inference is introduced in our procedure in order to make an optimal selection of statistical parameters in terms of the probability characteristics of the data. Therefore, in our procedure, we are free from any arbitrary reconstruction of the volume emission rates. In this paper, our method is compared with two conventional methods by applying them to the analysis of the rocket airglow data obtained by our group.

2. METHOD

Here we only describe the outline of the method, since its detail can be referred to the previous publication (Higuchi *et al.*, 1988).

* Present address: The Institute of Mathematical Statistics, Tokyo, Japan.

Representing a series of observed vertical column emission rate values as $\mathbf{I} = [I_1, I_2, \ldots]$, we can consider the observed data I_j to consist of three components as $I_j = T_j + S_j + \varepsilon_j$, where T_j, S_j and ε_j denote trend, periodic and random noise components, respectively. Our purpose is to find the most likely model of T_j and S_j from the observed data I_j.

We assume from our *a priori* knowledge that the trend component shows a locally smooth variation, and that the periodic component shows a locally sinusoidal variation. We adopt the constraints on the systematic components T_j and S_j, which represent these assumptions and are given by minimizing the second-order difference expressions $\Sigma[T_{j+1} - 2T_j + T_{j-1}]^2$ and $\Sigma[S_{j+1} - 2cS_j + S_{j-1}]^2$, respectively, where c is a constant determined by the characteristic period of the sinusoidal variation. The balance between the constraints for the systematic components S_j and T_j is controlled by a hyper-parameter s, which is a statistical parameter and cannot be determined by any *a priori* knowledge. The behaviour of periodic components comes to be more strictly sinusoidal with an increase of s. Furthermore, we assume that the systematic part $T_j + S_j$ does not deviate very largely from the observed data I_j. This leads to a constraint on the random component, which is given by minimizing square of the residuals $\Sigma\varepsilon_j^2$. We can control the balance between the constraints for the systematic components (T_j and S_j) and the random component (ε_j) by introducing a hyper-parameter d. As d increases, the constraint on both the systematic components becomes stronger, and then the residuals become larger. Note that the best model of T_j and S_j is determined by using the constrained least squares method for given s and d values.

We can determine the hyper-parameters s and d on the basis of the Bayesian statistics (Akaike, 1980). We assume the normal distribution of the second-order difference expressions of the three components, T_j, S_j and ε_j. For given data I_j, the optimal set of hyper-parameters (d, s) is what maximizes $L(d, s, \mathbf{I})$, the mean of the probability distributions of T_j, S_j and ε_j. Using $\mathrm{ABIC}(d, s) = -2 \log L(d, s, \mathbf{I})$ in place of $L(\cdot)$, where $\mathrm{ABIC}(d, s)$ stands for a Bayesian information criterion, the hyper-parameters s and d giving the minimum of ABIC should be optimal.

We have applied our method to the airglow data obtained by two rocket observations to compare it with two other conventional methods, the Fourier filtering method and the incremental straight line fitting method, which were suggested by Murtagh *et al.* (1984) to be available for deriving volume emission rates.

In applying the Fourier filtering method, we make the Fourier transformation of the data. Higher frequency components which are dominated by the random noise, are removed by multiplication with a filter function and then the inverse transformation is performed. In applying the incremental straight line fitting method, the data is differentiated to derive volume emission rates by least square fitting with a straight regression line in each altitude interval. The number of data points included in the altitude interval corresponds to a half of the number of data points included in the cut-off wavelength of a low-pass filter.

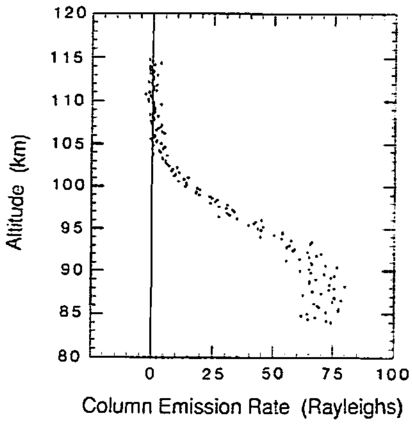

FIG. 1. VERTICAL COLUMN EMISSION RATES OF THE OI 557.7 nm NIGHT AIRGLOW IN UNITS OF RAYLEIGH (1 RAYLEIGH = 10^6 PHOTONS s^{-1} cm^{-2}-COLUMN) OBTAINED BY A ROCKET EXPERIMENT.

3. RESULTS AND DISCUSSION

Figure 1 shows the vertical column emission rates of the atomic oxygen airglow at 557.7 nm obtained by the rocket measurement over Uchinoura (31°N) at 13:00 U.T. on 6 September 1986 (Kita *et al.*, 1988b). The data are sampled every 0.2 km interval, and the number of the data sample is 156. The data exhibit a modulation with the same period as the rocket spin.

We applied our procedure to the data and obtained the most likely solution for trend, periodic and random components as shown on Fig. 2a. The hyper-parameters which control the balance among the constraints on these three components are determined to make the ABIC value minimum. We can see in Fig. 2a that the modulation in the observed data is estimated and eliminated satisfactorily, and consequently this procedure enables us to discuss the airglow features with a better altitude resolution. Defining the altitude resolution by the cut-off wavelength of the equivalent low-pass filter (we can consider smoothing and/or curve-fitting methods to be equivalent to low-pass filter methods), we have the altitude resolution

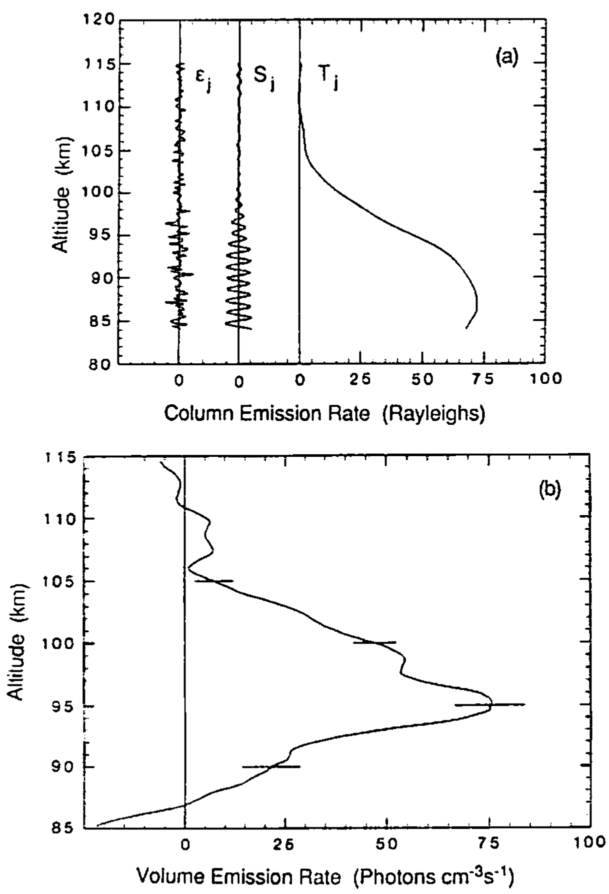

FIG. 2. (a) THE MODEL FOR THE DATA OF FIG. 1 AS INFERRED BY USING OUR METHOD AT THE MINIMUM OF ABIC.
T_j represents the estimation of the column emission rate, and S_j is the estimation of periodic modulation. ε_j shows the residual between $T_j + S_j$ and the observed data I_j. (b) Volume emission rates for the data of Fig. 1 as estimated by our inversion procedure.

$\delta z = 4.5$ km in this case. The values of the trend components T_j decrease rapidly in the altitude region 90–100 km, indicating the altitude distribution of the airglow layer. The increase of T_j values in the altitudes below 86 km can be explained by temporal and/or horizontal inhomogeneity of the airglow. Since the trend component is smaller than the amplitude of the random fluctuation ε_j above 105 km, we should be careful in discussing airglow features in this region. The amplitude of the periodic modulation S_j is found to correlate to the value of the trend component. This correlation suggests that the periodic modulation in the column emission rates is due to the horizontal inhomogeneity of the airglow. The amplitude of the random fluctuation component is larger below 95 km than above 100 km. We suppose that the temporal fluctuation in the photon flux of the airglow makes the random fluctuation larger. Since the random fluctuation is large (its amplitude is 5–10 R where 1 R = 10^6 photons s^{-1} cm^{-2}-column and its characteristic wavelength is 0.5–2 km), it is difficult to discuss small-scale structures in the airglow.

Figure 2b represents the most likely profile of the volume emission rate deduced in our method. In discussing any structures of the airglow emission layer, it is crucial to estimate errors in the inferred volume emission rates. However, the error estimation has a difficulty in almost all methods including ours. As mentioned above, the altitude resolution δz can be defined by the cut-off wavelength of the equivalent low-pass filter. The altitude variation of the trend component is considered to be linear within the altitude interval of δz. Therefore, we propose $\sqrt{2}\sigma/\delta z$ as an expression of the error in the estimated volume emission rate, where σ is the standard deviation of ε_j within the altitude range δz. Small-scale structures at altitudes of 91 and 97 km in the volume emission rate profile of Fig. 2b are not significant in consideration of both the error and the altitude resolution.

In Figs 3a, b and c, we compare the altitude profiles of volume emission rates derived from the column emission rates of Fig. 1 by using our method, the Fourier filtering method and the incremental straight line fitting method. The altitude resolution, which is defined by the cut-off wavelength, of the three methods is almost identical: about 5, 7.5 and 10 km in Figs 3a, b and c, respectively. The profile derived by using our method in Fig. 3a is identical to that of Fig. 2b. In these column emission rate data, the period and amplitude of the periodic modulation are not so large in comparison with those of the random fluctuation, and accordingly the conventional methods can reconstruct the volume emission rate models being consistent with our *a priori* knowledge. We see in Fig. 3b that the Fourier filtering method gives a model similar to our model shown on Fig. 3a. The incremental straight line fitting method, neglecting meaningless noises, also gives a model as on Fig. 3c similar to ours. However, since these conventional methods need to smooth out the periodic modulation, the altitude resolution is worse in these methods than in our method. Moreover, the most likely model cannot be specified within the frame of these conventional methods because of no criterion on the desirable profile. In our method, we can compare candidate models on the basis of Bayesian statistics. The ABIC value is about 840, 842 and 873 for the models on Figs 3a, b and c, respectively. As a consequence, we can designate the most likely model as shown on Fig. 3a.

Figure 4 shows the vertical column emission rate data of the NO γ (1,0) 214.8 nm band day airglow (Kita *et al.*, 1988a). The data are sampled every 0.5 km interval, and the number of the data sample is 218 in the rocket ascent and 224 in the rocket descent. The

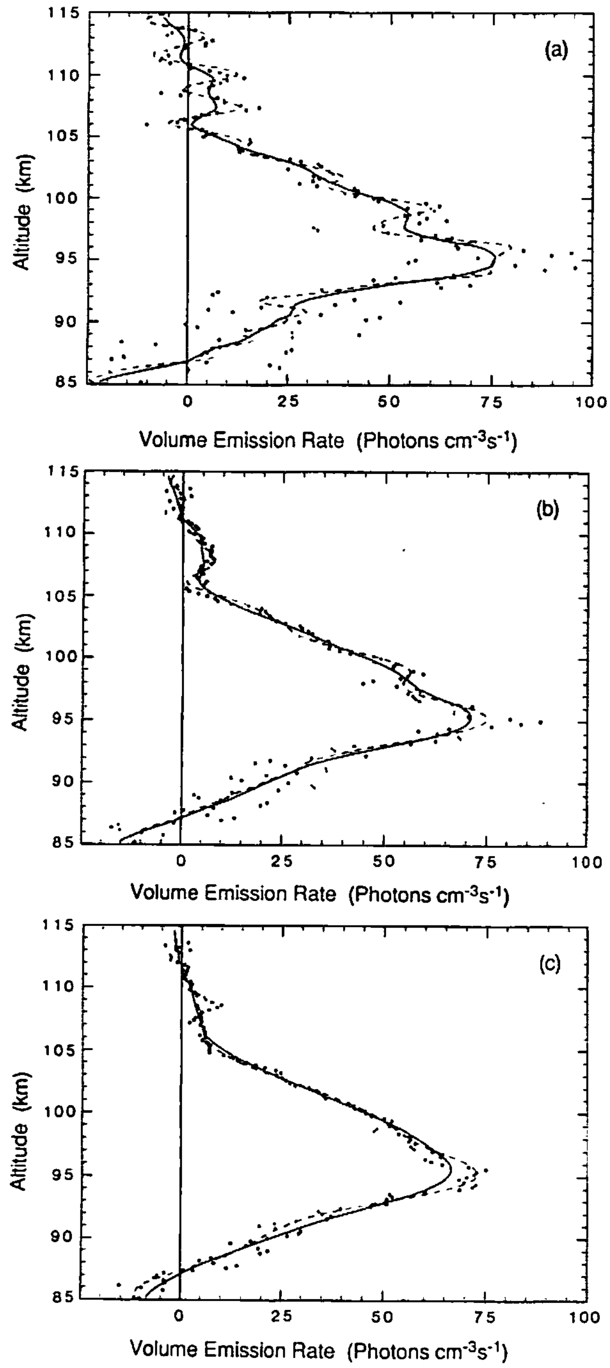

FIG. 3. VOLUME EMISSION RATES DEDUCED FROM THE DATA OF FIG. 1 BY USING OUR METHOD (SOLID CURVE), THE FOURIER FILTERING METHOD (DASHED CURVE) AND THE INCREMENTAL STRAIGHT LINE FITTING METHOD (DOTTED CURVE).
(a) Altitude resolution of 5 km; (b) altitude resolution of 7.5 km; and (c) altitude resolution of 10 km.

data suffered from a modulation probably caused by other experiments through electronic circuits, and the period and amplitude of this modulation is too large to infer the volume emission rate by using conventional methods without modeling of the periodic modulation. Applying the Fourier filtering method,

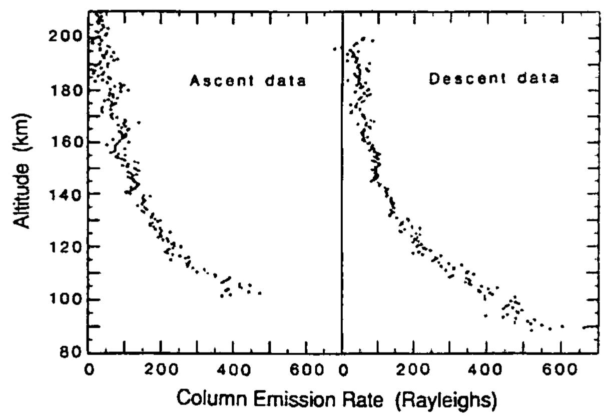

FIG. 4. VERTICAL COLUMN EMISSION RATES OF THE NO $\gamma(1, 0)$ 214.8 nm BAND DAY AIRGLOW OBTAINED BY A ROCKET EXPERIMENT.

we have obtained the altitude profile of NO number density which is proportional to volume emission rate and shown on Fig. 5. It is evident that this method is unable to eliminate the periodic modulation without distortion of the profile. We have tried to make a correction to reduce Fourier coefficients which correspond to the modulation period, and obtained the altitude profile also shown on Fig. 5 by the dotted curve. However, since the modulation consists of multiple Fourier components and the modulation component cannot be discriminated from the emission rate component, we cannot eliminate the modulation satisfactorily. On the contrary, we can eliminate the modulation by our procedure as shown in Fig. 6b, where the inferred model of three components is shown in Fig. 6a. Here we can reconstruct an airglow emission rate profile without degrading the altitude resolution. The altitude resolution is 27 km in our model. Figure 6a exhibits the amplitude of periodic

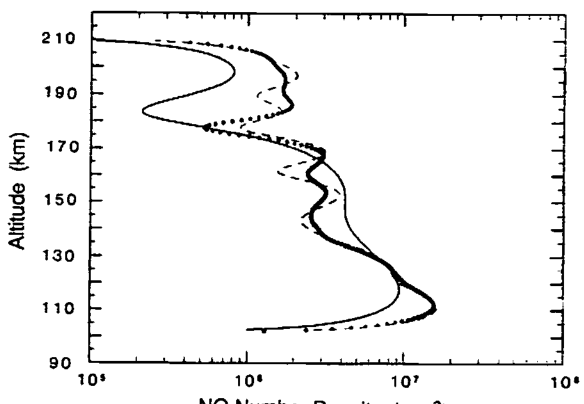

FIG. 5. NO NUMBER DENSITY DEDUCED FROM THE ROCKET ASCENT DATA OF FIG. 4 BY USING THE FOURIER FILTERING METHOD.
The altitude resolution is 72 km (solid curve) and 36 km (dashed curve and dotted line). The dotted curve shows the corrected volume emission rates.

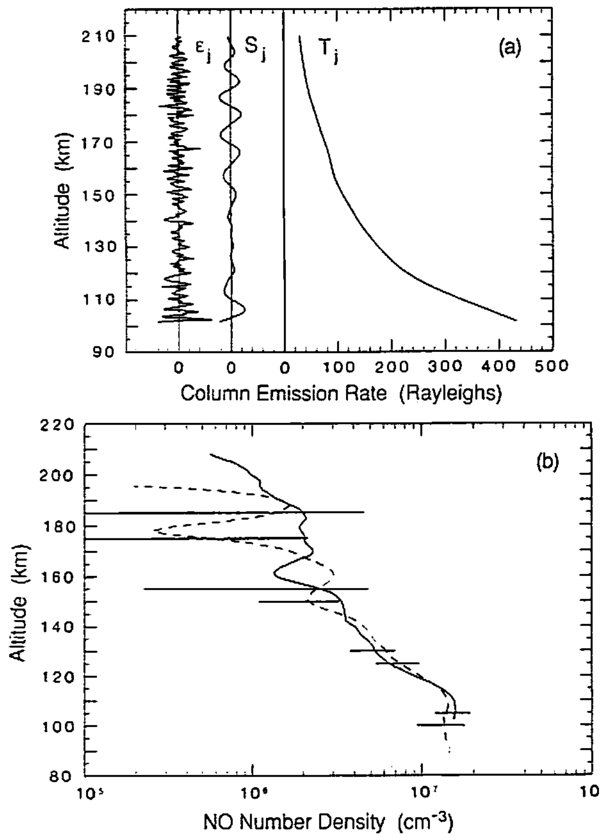

FIG. 6. (a) THE MODEL FOR THE ASCENT DATA OF FIG. 4 AS INFERRED BY USING OUR METHOD AT THE MINIMUM OF ABIC. T_j represents the estimation of the column emission rate, and S_j is the estimation of periodic modulation. ε_j shows the residual between T_j+S_j and observed data I_j. (b) NO number density deduced from the data of Fig. 4 using our inference procedure. The solid curve shows the density deduced from the rocket ascent data and the dashed curve shows that deduced from the rocket descent data.

component S_j varying with height in accordance with the period of the rocket precession. This variation makes it more difficult to eliminate the periodic modulation. Even in our model, trend and random components are not completely free from the influence of the periodic modulation. We see wavy structures reflecting this influence in the altitude profile of NO number density above 140 km in Fig. 6b. The profile of NO number density normally shows a peak at a height of around 110 km. However, the altitude resolution is not good enough to make the peak appear clearly. Moreover, since this altitude region is near the end of data set, we suppose that the periodic component is not well separated from the trend component in this region.

4. SUMMARY

We have demonstrated that the Bayesian procedure is effective in inferring the most likely volume emission rates from the observed data in comparison with two conventional methods, the Fourier filtering method and the incremental straight line fitting method. Our method has two major advantages in comparison with conventional methods. First, our method has a statistical criterion to determine the optimal model of the emission rate profile. The models selected according to the criterion are consistent with our *a priori* knowledge. In conventional methods, we are are not free from arbitrary selection of the model, and this lack of criterion leads us to sterile discussion on airglow structures. Secondly, in our present method, we can infer and eliminate a troublesome periodic modulation. Therefore, when observed data are contaminated by a spurious modulation, our method can obtain an emission rate profile with better altitude resolution and less distortion than conventional methods.

We have adopted the constraint of minimizing the second-order difference for trend components, and assumed a locally sinusoidal variation for periodic components. This procedure is useful and effective enough to analyze the atomic oxygen green line at 557.7 nm and NO $\gamma(1, 0)$ band airglow data obtained by rocket observations. It is also possible to adopt another constraint in the model according to *a priori* information on the trend and periodic components in the observed data, or to add another component in the model.

REFERENCES

Akaike, H. (1980) Likelihood and the Bayes procedure, in *Bayesian Statistics* (Edited by Bernardo, J. M., De Groot, M. H., Lindley, D. V. and Smith, A. F. M.). University Press, Valencia, Spain.

Higuchi, T., Kita, K. and Ogawa, T. (1988) Bayesian statistical inference to remove periodic noises in the optical observation aboard a spacecraft. *Appl. Optics* **27**, 4514.

Kita, K., Iwagami, N. and Ogawa, T. (1988a) Thermospheric NO measurement at solar minimum. *J. Geomagn. Geoelect.* **40**, 261.

Kita, K., Iwagami, N., Ogawa, T., Miyashita, A. and Tanabe, H. (1988b) Height distributions of the night airglow emissions in the O_2 Herzberg I system and oxygen green line from a simultaneous rocket observation. *J. Geomagn. Geoelect.* 40, 1097.

Murtagh, D. P., Greer, R. G. H., McDade, I. C., Llewellyn, E. J. and Bantle, M. (1984) Representative volume emission profiles from rocket photometer data. *Ann. Geophysicae* **2**, 467.