Improvement in the technique to extract gravity wave parameters from lidar data

Guotao Yang,[1] Barclay Clemesha,[1] Paulo Batista,[1] and Dale Simonich[1]

Received 3 October 2007; revised 22 June 2008; accepted 2 July 2008; published 10 October 2008.

[1] From the analysis of 11 years of INPE lidar profiles and some data from the WIPM lidar, we find that the technique proposed by Gardner and Voelz (1987) for gravity wave parameter extraction does not work well at these sites. At São José dos Campos (23° S) and Wuhan (31°N), Na_s occurrence is very high, and the background layer is often far from a single symmetrical Gaussian, with the result that most spatial power spectra (~80%) will be distorted as the given examples show. The power spectrum is used to extract wave parameters in the Gardner and Voelz (1987) technique, but this only works well for an undistorted spectrum. To solve this problem, a new method is developed. First, a "connection layer" is constructed, allowing the determination of approximate wave parameters and an improved background layer. The background layer parameters and the wave parameters are then used as the initial input parameters for a Levenberg-Marquardt fit, leading to a much better fitted background layer and better wave parameters. This new method achieves good agreement between simulated sodium variations and observations, as well as artificial numerically simulated "observations," as shown in the examples presented in this paper.

Citation: Yang, G., B. Clemesha, P. Batista, and D. Simonich (2008), Improvement in the technique to extract gravity wave parameters from lidar data, *J. Geophys. Res.*, *113*, D19111, doi:10.1029/2007JD009454.

1. Introduction

[2] Gravity waves have great influence on the circulation, structure and composition of the middle and upper atmosphere. Observations of gravity wave activity have focused on characterization of continuous gravity wave spectra and measurements of quasi-monochromatic wave parameters. Measurement of quasi-monochromatic wave parameters can provide information on the characteristics of the gravity waves, such as their distributions and the relationships between their parameters.

[3] Sodium lidar is a useful tool for studying gravity waves in the mesopause region, as it can measure the gravity wave perturbations in the vertical direction with a resolution of a few hundred meters. Moreover, the sodium layer is a good tracer of gravity wave perturbations because the wave amplitudes are usually large near the mesopause and the steep sodium density gradients on the bottom and topsides of the layer enhance the wave perturbations. Most lidar studies of gravity waves have concentrated on quasi-monochromatic waves or wave events, as the dominant wave perturbations in lidar data tend to have a monochromatic appearance. The first report of wavelike structures in the sodium layer was made by *Rowlett et al.* [1978]. Subsequently, many authors reported the observation of oscillations attributable to gravity waves in sodium lidar data [e.g., *Shelton et al.*, 1980; *Chanin and Hauchecorne*, 1981; *Kamiyama and Tomita*, 1985].

[4] Gardner and his colleagues at Illinois have developed a fundamental theory to describe gravity wave perturbations in passive chemical tracers [*Gardner and Shelton*, 1985; *Gardner and Voelz*, 1985, 1987]. *Gardner and Voelz* [1985] simplified their analysis of monochromatic waves by assuming the unperturbed layer could be described by a Gaussian function, and proposed a technique for the quantitative estimation of quasi-monochromatic wave parameters. *Gardner and Voelz* [1987] (hereinafter referred to as GV87) perfected this technique and made a detailed analysis of monochromatic waves observed at Urbana, Illinois (40°N, 88°W). In the GV87 technique, the wave parameters are extracted from the spatial power spectrum and the background sodium layer is assumed to have an approximately Gaussian shape. Applying this technique, monochromatic wave parameters have been obtained from sodium lidar data at a number of different latitudes [*Beatty et al.*, 1992; *Collins et al.*, 1994].

[5] The analysis technique developed by the Illinois workers appears to be adequate in the absence of sporadic (Na_s layers), but *Clemesha et al.* [1998a] showed that the occurrence of Na_s layers close to the peak of a propagating wave in the sodium layer seems to be a common feature of the lidar data from São José dos Campos (23°S, 46°W). The analysis of the wave structures observed in the sodium layer by *Clemesha et al.* [1998b] showed that the associated oscillations in the sodium mixing ratio are frequently too

[1]Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brazil.

0148-0227/08/2007JD009454

large to be ascribed to simple gravity wave perturbation. A report on the simultaneous measurement of the sodium layer and meteor winds made by *Clemesha et al.* [2001], reinforced their previous suggestion that the mechanism responsible for the formation of Na_s also generates at least some of the more modest wave modulation of the sodium layer, often seen to be associated with Na_s layers. *Simonich et al.* [2005] presented long-term average sodium profiles for data with and without the presence of sporadic layers, and showed that the presence of Na_s has negligible effect on such long-term averages, suggesting that at least part of the normal sodium layer originates in a source that also gives rise to Na_s, which would occur when the source itself becomes layered. These studies indicate that Na_s often accompanies wave perturbed layers, and that the source of the background sodium layer may be variable.

[6] In this paper, from our many years of lidar data at São José dos Campos (23°S, 46°W) and some data from Wuhan (31°N), China, we find that weak Na_s and irregular background layer shape can have a large effect on the wave analysis, and the technique proposed by GV87 for gravity wave parameter extraction does not work well under these conditions. Sometimes even large errors can ensue if the GV87 technique is applied. Ambiguous Na_s layers and irregular background layer shape are very common in our lidar data. In view of this problem, we developed a new method for extracting the wave parameters. It should be noted that this work does not invalidate that of *Gardner and Voelz* [1985, 1987] when applied to observations with a well behaved background layer, but the new method, also based upon the fundamental theory developed by Gardner and his colleagues, does give much better results when the background layer is far from Gaussian or in the presence of sporadic layers, which is frequently the case at our location.

[7] The rest of this paper is divided up as follows: After an introduction to the GV87 technique and a description of the lidar data, we present the methodology of our new method. For better description for each step of the methodology, we also make some numerical simulations of wave perturbation layers with artificial parameters, and demonstrate how to deal with these simulated layers in each step. Then we show some observation examples and extract wave parameters from them, using the previously introduced methodology. The following sections are the power spectrum statistics and the discussion. Finally, we present our conclusions.

2. Technique for Extracting Gravity Wave Parameters Used in GV87

[8] Assuming the steady state profile of the sodium layer is horizontally homogeneous, GV87 gave the layer response to a monochromatic gravity wave:

$$n_s\left(\vec{r},t\right) = \frac{n_0\left(z-\gamma H\ln\left[1+\left[Ae^{\beta z}/(\gamma-1)\right]\cos\left(\omega t-\vec{k}\cdot\vec{r}\right)\right]\right)}{\left\{1+\left[Ae^{\beta z}/(\gamma-1)\right]\cos\left(\omega t-\vec{k}\cdot\vec{r}\right)\right\}} \quad (1)$$

Where: $n_s(\vec{r}, t)$ is the sodium layer response to the gravity wave; $n_0(z)$ is the unperturbed sodium layer; $Ae^{\beta z}$ is the wave amplitude; β is the amplitude growth factor; ω is the wave frequency; $\vec{r} = x \cdot \hat{x} + z \cdot \hat{z}$, is the position vector where x is the horizontal coordinate and z is the vertical coordinate. $\vec{k} = k_x \cdot \hat{x} + k_z \cdot \hat{z}$, is the wave number vector; γ is the ratio of specific heats ($\sim$1.4); H is the atmospheric scale height ($\sim$6 km);

[9] Considering only the linear perturbations, the magnitude of the positive frequency component of the time averaged spatial power spectrum is given by:

$$\begin{aligned}\overline{|\Phi_s(k)|^2} = \frac{1}{\tau}\int_{t_0-\tau/2}^{t_0+\tau/2}|\Phi_s(k)|^2dt &= \left[\frac{\gamma HA}{2(\gamma-1)}\right]^2\cdot\left[\left(\frac{1}{\gamma H}-\beta\right)^2 + (k-k_z)^2\right]|\Phi_0(k-k_z-i\beta)|^2 \\ &+ |\Phi_0(k)|^2 - \frac{\gamma HA}{2(\gamma-1)}\frac{\sin(\omega\tau/2)}{\omega\tau/2} \\ &\cdot \mathrm{Re}\left\{\left[\left(\frac{1}{\gamma H}-\beta\right)+i(k-k_z)\right] \cdot \Phi_0^*(k)\Phi_0(k-k_z-i\beta)e^{-i\omega t_0}\right\} \end{aligned} \quad (2)$$

Near k_z, the last two items on the right of (2) can be neglected when k_z is not too small provided we assume the unperturbed sodium layer to be a Gaussian:

$$n_0(z) = \frac{c_0}{\sqrt{2\pi}\sigma_0}e^{-(z-z0)^2/2\sigma_0^2} \quad (3)$$

here, c_0 is the layer column abundance; z_0 is the center height; σ_0 is the RMS width;

[10] The criterion to neglect the last two items on the right of (2) is given by GV87:

$$\exp\left(-k_z^2\sigma_0^2/2\right) \ll Ae^{\beta z_0}/2(\gamma-1) \quad (4)$$

[11] *Gibson-Wilde et al.* [1996] studied the effect of the sodium layer width on the GV87 technique by using a numerical model, and found [*Gibson-Wilde et al.*, 1996, p. 9522] "that quasi-monochromatic gravity waves with vertical wavelengths larger than approximately 10 km may not be reliably retrieved from Na lidar data," so the critical value of k_z is about $2\pi/10$ km.

[12] The last two items on the right of (2) are neglected, and only the first item makes a contribution, with two maxima and a minimum at the spatial frequency $k = k_z$. This symmetrical structure is the characteristic of the gravity wave spectral signature as shown by Figure 2 or Figure B1 of GV87. Using only the magnitudes of the maxima and minimum, the amplitude growth factor β and wave amplitude $Ae^{\beta z}$ can be conveniently obtained.

[13] In the GV87 technique, the gravity wave parameters are extracted from the power spectrum, and only the spectra having symmetrical characteristic as in Figure B1 of GV87 are considered. This technique is based on fitting a single Gaussian, so it works when the background layer is an approximate Gaussian. However, when the background layer is far from a normal Gaussian, this technique does

not work, as the numerical simulations presented in section 4.3 show.

3. Observational Data

[14] Our sodium lidar measurement technique and the precision of measurement have been described in earlier papers [e.g., *Simonich et al.*, 1979; *Clemesha et al.*, 1992]. This sodium lidar has been used to observe the sodium layer for more than 30 years, but the height resolution was only 1 km before 1993, too large for shortwave analysis. The lidar height resolution for data obtained after 1994 is 250 m or 300 m, and the time resolution is 3 min or 5 min. The photon count at the peak of the sodium layer is about 300, and the error in wave parameters extraction from photon noise is typically about 15% or less, which is estimated from *Gardner and Voelz* [1987] and *Senft and Gardner* [1991]. The height resolution of the Wuhan Institute of Physics and Mathematics (WIPM) lidar is 200 m, and time resolution is 3 min. The sodium density profiles are first spatially and temporally low-pass filtered with cutoffs of about 2 km and 30 min.

4. Methodology of the New Technique

4.1. Theoretical Consideration of the Power Spectrum Distorted by an Irregular Background Layer

[15] We will first discuss the effect of the irregular background layer shape on the wave parameter extraction from theoretical consideration of the power spectrum.

[16] For the simplest case, the background layer is an asymmetrical Gaussian shape:

$$n_0(z) = \begin{cases} c \cdot e^{-(z-z_0)^2/2\sigma_1^2} & z < z_0 \\ c \cdot e^{-(z-z_0)^2/2\sigma_2^2} & z \geq z_0 \end{cases} \tag{5}$$

$$\begin{aligned} \Phi_0(k) &= \int_{-\infty}^{\infty} n_0(r,t) e^{ikz} dz \\ &= \int_{z_0-L/2}^{z_0} c \cdot \mathrm{e}^{-\frac{(z-z_0)^2}{2\sigma_1^2}} e^{ikz} dz + \int_{z_0}^{z_0+L/2} c \cdot \mathrm{e}^{-\frac{(z-z_0)^2}{2\sigma_2^2}} e^{ikz} dz \\ &\approx \sqrt{\frac{\pi}{2}}/2 \ \sigma_1 \cdot c \cdot \exp\left(\frac{-k^2\sigma_1^2}{2}\right) \cdot \exp(ikz_0) + \sqrt{\frac{\pi}{2}}/2 \ \cdot \sigma_2 c \\ &\quad \cdot \exp\left(\frac{-k^2\sigma_2^2}{2}\right) \cdot \exp(ikz_0) \\ &= \frac{\Phi_1(k)}{2} + \frac{\Phi_2(k)}{2} \end{aligned} \tag{6}$$

[17] The power spectrum is still obtained under formula (2), and the second item on the right hand side of formula (2) is:

$$|\Phi_0(k)|^2 \approx \frac{|\Phi_1(k)|^2}{4} + \frac{|\Phi_2(k)|^2}{4} + \mathrm{Re}\left[\frac{\Phi_1^*(k)\Phi_2(k)}{2}\right] \tag{7}$$

here, $|\Phi_2(k)|^2$ is the spectrum corresponding to the part of the layer above the height of peak Na density, and if its RMS width is too small, this item will extend into the frequency region where only the wave should make a contribution. So the characteristic of the gravity wave spectral signature will be distorted, which is the case of a single Gaussian layer distorting the power spectrum of a long-wavelength wave, as pointed out by *Gibson-Wilde et al.* [1996]. A similar distortion factor will also occur in the first item of equation (2). If $\sigma_2 \sim 2$ km, the characteristic spectral signature (symmetrical shape with two maxima and a minimum) of a wave ($\lambda_z \geq 4$ km) will be distorted. So, if the background layer has a steep gradient, the characteristic spectral signature will be distorted, and the GV87 technique will not be suitable.

[18] There is also another kind of background layer shape that can distort the characteristic spectral signature of the wave. In our data, the background sodium layer sometimes is very wide, something like a double Gaussian. In a simple approximation, the background layer can be written as:

$$n_0(z) = c \cdot e^{-(z-z_1)^2/2\sigma^2} + c \cdot e^{-(z-z_2)^2/2\sigma^2} \tag{8}$$

$$\text{so}: |\Phi_0(k)|^2 = \pi\sigma^2 c^2 \cdot \exp\left(-k^2\sigma^2\right) \cdot \{1 + \cos[k(z_1 - z_2)]\} \tag{9}$$

if we use: $k_t = k - k_z$ and assume $e^{\beta z1} \approx e^{\beta z2}$, we obtain:

$$\begin{aligned} &|\Phi_0(k - k_z - i\beta)|^2 \\ &\quad \approx \pi\sigma^2 c^2 \cdot \exp\left[-\left(k_t^2 - \beta^2\right)\sigma^2\right] \cdot e^{2\beta z_1} \cdot \{1 + \cos[k_t \cdot (z_1 - z_2)]\} \end{aligned} \tag{10}$$

The region of k_t is about $(-0.2, 0.2)$, so the range of variation of $\cos[k_t \cdot (z_1 - z_2)]$ is (1, 0.7) when $z_1 - z_2 = 4$] km, and (1, −0.03) when $z_1 - z_2 = 8$ km. So $|\Phi_0(k - k_z - i\beta)|^2$ will change greatly when $z_1 - z_2$ is large, and the characteristic spectral signature will also be distorted.

4.2. Steps of the New Methodology

[19] We first give here the steps of the methodology:

[20] 1. Connect the midpoints between each peak and valley of the layer to construct an approximate background layer (hereafter, connection layer).

[21] 2. From the connection layer, an approximate linear wave perturbation of the atmospheric density is derived, which also can be used to recognize minor Na_s.

[22] 3. The previously obtained wave perturbation is high-pass filtered to derive approximate wave parameters.

[23] 4. From the approximate wave parameters, an improved background layer can be obtained.

[24] 5. Several Gaussians are used to approximate the improved background layer.

[25] 6. The Gaussians background layer in step 5 and the wave parameters in step 3 are used as the initial input parameters for Levenberg-Marquardt calculations, which are based on formula (1). After performing Levenberg-Marquardt calculations, the final wave parameters and background layer are obtained.

4.3. Numerical Simulations of Wave Perturbation Layers With Artificial Parameters

[26] From formula (1), we can create several wave perturbation layers with artificial parameters in order to test

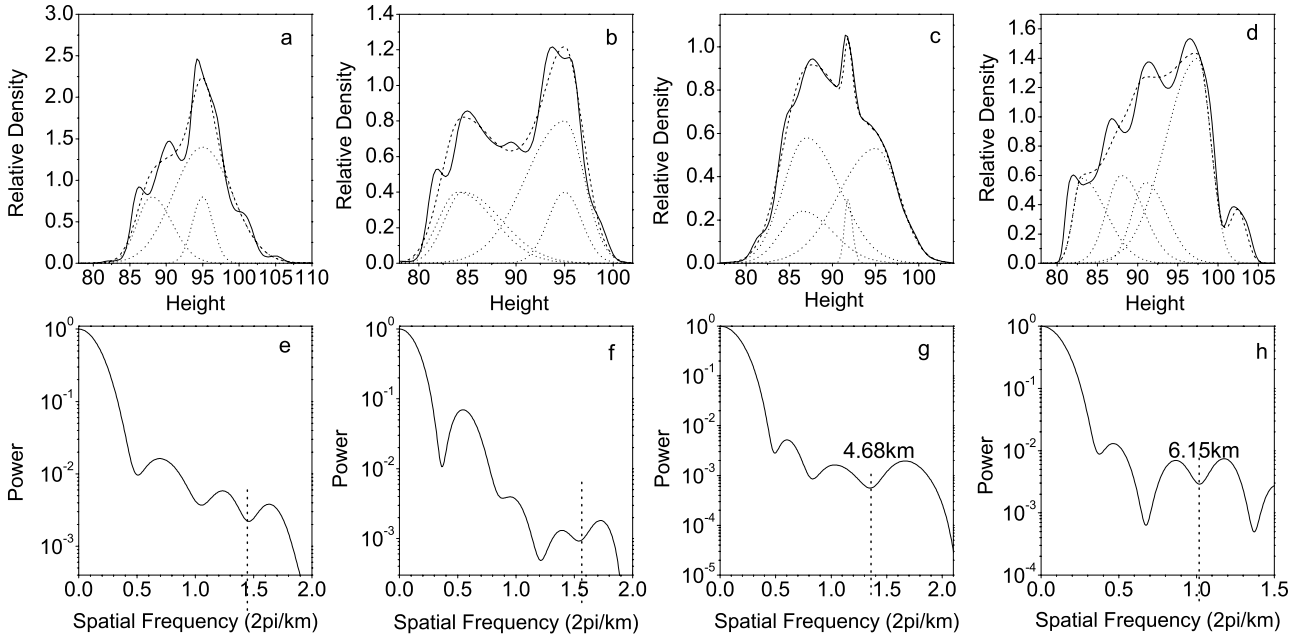

Figure 1. (a) The numerical simulated wave perturbation layer (“observed layer,” solid curve) is based on the assumed background layer (dashed curve), which is composed by three Gaussians (dotted curves) for case 1. (b) The same as Figure 1a but for case 2. (c) The same as Figure 1a but for case 3. (d) The same as Figure 1a but for case 4. (e–h) The normalized power spectra for the “observed layers” in Figures 1a–1d, respectively.

the new technique. Four cases are given in Figure 1. The assumed background layers are shown by the dotted curves in Figures 1a–1d. With these assumed background layers (their parameters are shown in Tables 1 and 2) and the assumed wave parameters (shown in Table 3), simulated layers can be calculated from formula (1), as the solid curves shown in Figures 1a–1d.

[27] The model background layer shown in Figure 1a consists of three unsymmetrical Gaussians (hereafter, Gaussian). The smallest Gaussian, whose peak is at 95 km, makes the top of the layer steep, while the lowest Gaussian makes the bottom steep. So the power spectrum (shown in Figure 1e) of the simulated layer (the solid curve in Figure 1a) does not have a symmetrical structure: There is a minimum near the assumed $k_z = 2\pi/4.4$ km, but the value of the two adjacent maxima is unequal. The maximum corresponding to a lower frequency is obviously higher than the one corresponding to a higher frequency. So for this case, the GV87 technique does not work, because this technique only considers a spectrum which fits the model illustrated in Figure B1 of GV87.

[28] The assumed background layer shown in Figure 1b consists of four Gaussians. Two Gaussians are at greater heights and two are at lower heights. The long distance (about 10 km) between these two groups of Gaussians also distorts the power spectrum (shown in Figure 1f): The value of the two maxima around $k_z \approx 2\pi/4$ km is not equal. So for this case, the GV87 technique also cannot be adopted.

[29] The assumed background layer shown in Figure 1c consists of four Gaussians, with a sharp little Gaussian at 92 km acting as Na_s. Here, we assume the Na_s layer is fixed in height but perturbed by the wave under formula (1). The long distance between the Gaussians and the existence of a sharp peak excessively distort the power spectrum: The minimum corresponding to the assumed $k_z = 2\pi/3.4$ km disappears, but a minimum corresponding to $k_z = 2\pi/4.7$ km is present. Although the adjacent maxima have equal values, the GV87 technique does not work well, as the wavelength obtained is much too long.

[30] The assumed background layer shown in Figure 1d consists of five Gaussians, with a small sharp Gaussian at 102.5 km acting as Na_s. The power spectrum is seriously distorted by this irregular background layer: The minimum corresponding to the assumed $k_z = 2\pi/5$ km disappears, but a minimum corresponding to $k_z = 2\pi/6.15$ km appears.

[31] Note that in the following discussion the numerical simulated layer is referred to as the “observed layer.”

4.4. Demonstration of How to Extract Wave Parameters Using the Sequence Given in Section 4.2

[32] The GV87 technique does not work well for the above “observed layers,” but accurate wave parameters can be extracted following the steps given in 4.2.

4.4.1. Step 1

[33] We can assume that the peaks and valleys of the “observed layers” correspond to the antinodes of the wave, while the midpoints between each peak and valley correspond to the nodes, where the density should be equal to that of the background layer. So we can connect these midpoints to construct an approximate background layer (connection layer). Here, the midpoints are the points where $d^2 n_s(z)/dz^2 = 0$. In this way, we obtain a reasonable background layer.

Table 1. Gaussians for Construction of the Background Layer of Case 1 and Case 2 in Section 4[a]

	Assumed	Obtained in Step 3	Obtained in Step 6
	Case 1		
Gaussian 1			
c	1.4	1.27	1.38
z_0 (km)	95	94.3	95.0
σ_1 (km)	4	3.8	4.13
σ_2 (km)	4	4.6	4.05
Gaussian 2			
c	0.8	0.90	0.81
z_0 (km)	95	95.0	95.0
σ_1 (km)	1.41	1.35	1.36
σ_2 (km)	1.41	2.07	1.49
Gaussian 3			
c	0.8	0.74	0.76
z_0 (km)	88	87.5	87.9
σ_1 (km)	2	1.99	1.96
σ_2 (km)	2.83	3.57	2.91
	Case 2		
Gaussian 1			
c	0.4	0.271	0.272
z_0 (km)	84	82.4	81.86
σ_1 (km)	2	1.53	1.26
σ_2 (km)	2.65	10.0	7.1
Gaussian 2			
c	0.4	0.53	0.58
z_0 (km)	84.5	85.5	84.4
σ_1 (km)	1.32	2.61	1.40
σ_2 (km)	4	4.50	5.61
Gaussian 3			
c	0.8	1.00	1.07
z_0 (km)	95	95.0	95.2
σ_1 (km)	4	2.46	2.58
σ_2 (km)	2	1.77	1.72
Gaussian 4			
c	0.4		
z_0 (km)	95		
σ_1 (km)	1.22		
σ_2 (km)	1.22		

[a]The form of the Gaussian is under formula 5. In step 3, we only give the Gaussians which are used as the initial background layer in step 6.

[34] The dotted curves shown in Figures 2a–2d are the obtained connection layers corresponding to the "observed layers" in Figures 1a–1d, respectively.

4.4.2. Step 2

[35] Once the sodium variation Δns has been obtained, the linear gravity wave perturbation of the atmospheric density ($\frac{\Delta N}{N} \simeq Ae^{\beta z}\cos(\omega t - \vec{k}\cdot\vec{r})$, hereafter, wave perturbation) can be obtained by the linear layer response [*Gardner and Shelton*, 1985, formula 41]:

$$\frac{\Delta n_s}{n_0} \simeq -\frac{1+\frac{\gamma H}{n_0}\frac{dn_0}{dz}}{\gamma-1}\frac{\Delta N}{N} \simeq -\frac{1+\frac{\gamma H}{n_0}\frac{dn_0}{dz}}{\gamma-1}Ae^{\beta z}\cos\left(\omega t - \vec{k}\cdot\vec{r}\right) \tag{11}$$

Here, $\Delta n_s = n_s - n_0$. From the formula, we can see that when $|(1 + \gamma H dn_0/dz/n_0)/(\gamma - 1)|$ is very small, the calculated wave perturbation will have a large error, so we only calculate the wave perturbation where the absolute value of the linear amplified factor is not less than about one. Generally, the heights where the linear wave perturbation vanishes are around 2 km higher than the layer peak. From formula (11), the wave perturbation can be calculated. Shown in Figures 2e–2h are the wave perturbations corresponding to the connection layers in Figures 2a–2d, respectively.

[36] Every antinode of the wave perturbation is enumerated in sequence, starting at the lowest height. In Figure 2g, we can see that there is a prominent antinode at 92 km, which we will refer to as antinode 8. The wave amplitude at this height is too large, resulting in $1/\beta$ being very small, from the average of antinodes 5 and 6 to the average of 7 and 8, about 2.9 km. Moreover, the distance between antinodes 8 and 6 is 1.5 km longer than other antinode interval distances. So we believe that this prominent anti-

Table 2. Gaussians for Construction of the Background Layer of Case 3 and Case 4 in Section 4[a]

	Assumed	Obtained in Step 3	Obtained in Step 6
	Case 3		
Gaussian 1			
c	0.24	0.827	0.851
z_0 (km)	86.5	86.9	86.9
σ_1 (km)	2.45	2.76	2.67
σ_2 (km)	3.54	3.87	3.74
Gaussian 2			
c	0.58	0.488	0.524
z_0 (km)	87	94.4	95.0
σ_1 (km)	2.7	2.76	3.85
σ_2 (km)	3.81	3.47	2.64
Gaussian 3			
c	0.53		0.31
z_0 (km)	95		91.7
σ_1 (km)	4.24		0.37
σ_2 (km)	2.65		0.54
Gaussian 4			
c	0.3		
z_0 (km)	91.8		
σ_1 (km)	0.4		
σ_2 (km)	0.5		
	Case 4		
Gaussian 1			
c	0.58	0.38	0.53
z_0 (km)	83	82.6	82.9
σ_1 (km)	1	1.18	1.35
σ_2 (km)	3.1	4.05	5.57
Gaussian 2			
c	0.6	0.95	0.97
z_0 (km)	88	90.8	90.5
σ_1 (km)	1.87	4.41	4.39
σ_2 (km)	2.74	2.31	3.96
Gaussian 3			
c	0.55	1.41	1.39
z_0 (km)	91	96.9	97.5
σ_1 (km)	1.73	3.24	4.95
σ_2 (km)	2.65	2.37	2.26
Gaussian 4			
c	1.4		3.59
z_0 (km)	97.5		102.5
σ_1 (km)	4		1.37
σ_1 (km)	1.58		1.43
Gaussian 5			
c	0.36		
z_0 (km)	102.5		
σ_1 (km)	1		
σ_2 (km)	1		

[a]The form of the Gaussian is under formula 5. In step 3, we only give the Gaussians which are used as the initial background layer in step 6.

Table 3. Gravity Wave Parameters Referred to in Section 4

	Assumed	Obtained In Step 3	Obtained In Step 6
	Case 1		
λ_z (km)	4.4	4.38	4.385
$Ae^{\beta*92km}$ (%)	4.0	3.89	3.96
β (km)	40	35.45	40.8
Error of amplitude (%)		2.75	1.0
	Case 2		
λ_z (km)	3.97	3.85	3.83
$Ae^{\beta*92km}$ (%)	2.4	2.60	2.36
β (km)	−36	−78	−35.5
Error of amplitude (%)		8.3%	1.67
	Case 3		
λ_z (km)	3.38	3.33	3.31
$Ae^{\beta*92km}$ (%)	1.0	1.29	0.923
β (km)	−25	−180.3	−19.6
Error of amplitude (%)		29	7.7
	Case 4		
λ_z (km)	5.02	4.96	5.00
$Ae^{\beta*92km}$ (%)	3.0%	2.64	3.034
β (km)	−33.3	−26.8	−32.49
Error of amplitude (%)		12	1.13

node, corresponding to the peak of the "observed layer" at 92 km, is just the effect of the Na_s.

[37] In Figure 2h, there is a prominent antinode near 102 km, which is also too large, and results in a very small $1/\beta$ (7.6 km). So this prominent antinode, which corresponds to the peak of the "observed layer" at 102 km, is also the effect of the Na_s. In Figures 2e and 2f, no prominent antinode corresponding to the "observed layer" peak is observed.

4.4.3. Step 3

[38] At heights where a Na_s is present, it is not easy to determine the background layer shape, so we simply discard the wave perturbation at these heights. The dotted curves in Figures 3a and 3b represent the same wave perturbations, as in Figures 2e and 2f, respectively. But the wave perturbations shown in Figures 3c and 3d do not include the oscillation at Na_s heights.

[39] From the obtained wave perturbation, the wavelength can be roughly estimated by the distance between the adjacent antinodes. We filter the wave perturbation with a high-pass filter whose cutoff is about 2 km greater than the estimated vertical wavelength. The filtered wave perturbation can be used to extracted wave parameters: The wavelength can be obtained by measuring the average distance between each adjacent antinode. The wave amplitude ($Aexp^{\beta*92km}$) and its growth factor (β) can be obtained by making a linear fit to these wave perturbation antinodes. The linear fit is:

$$\log(Y_i) = \log(A) + \beta \cdot z_i \tag{12}$$

Here, Y_i is the absolute value for each antinode; z_i is the height for each antinode;

[40] The solid curves shown in Figure 3 are the high-pass filtered wave perturbations while the dashed curves are the discarded lower-frequency components. From Figure 3, we can see that the filtered wave perturbations are more regular. The extracted wave parameters are listed in Table 3. We can see that the error of the obtained wave amplitude in this step is about 10–30%.

4.4.4. Step 4 (Create an Improved Background Layer)

[41] First, we only consider the background layer shape at heights between two adjacent wave antinodes. For example, we consider how to construct the background layer between

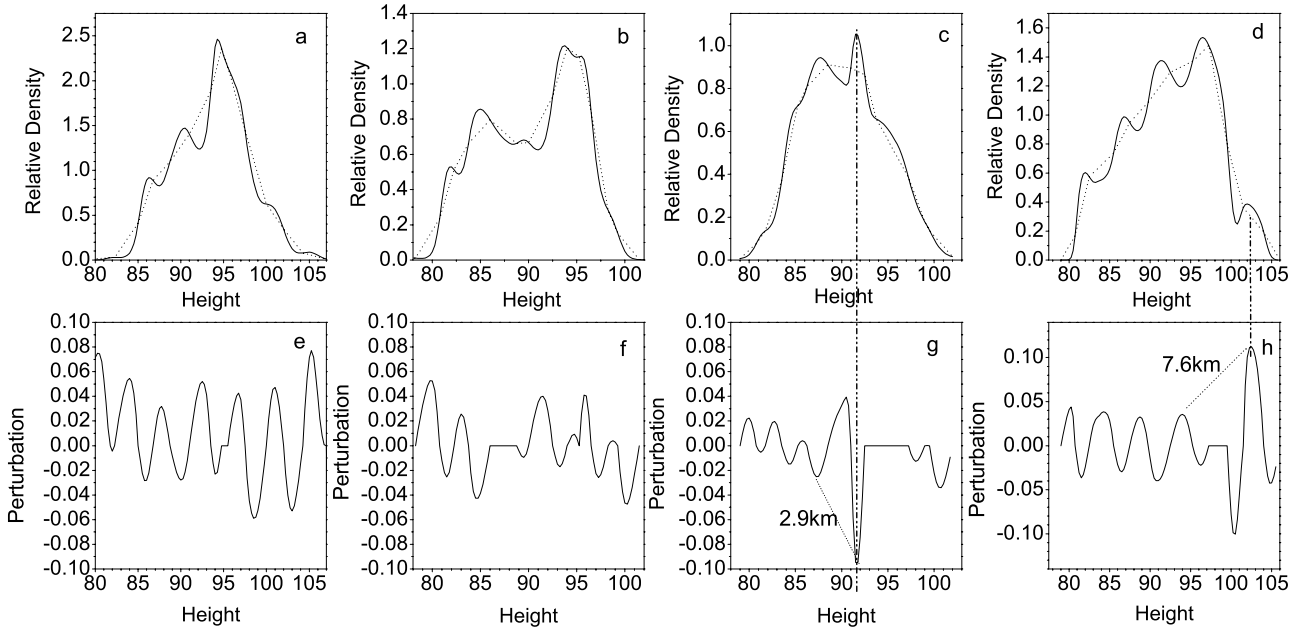

Figure 2. The "observed layer" (solid curve) and the connection layer (dotted curve) for (a) case 1, (b) case 2, (c) case 3, and (d) case 4. The wave perturbation (solid curve) based on the "observed layer" and the connection layer for (e) case 1, (f) case 2, (g) case 3, and (h) case 4.

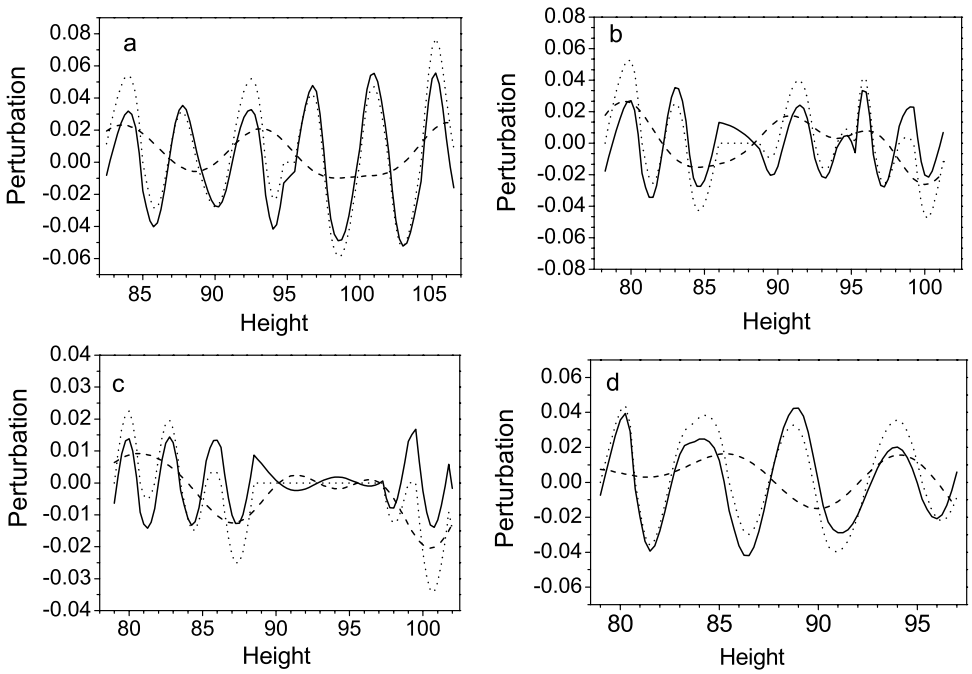

Figure 3. (a) The wave perturbation (dotted curve) is high-pass filtered to produce filtered wave perturbation (solid curve) for case 1. The cutoff is 6 km and the dashed curve is the discarded lower-frequency part. (b–d) As for Figure 3a but for cases 2, 3, and 4, with cutoffs of 5.5, 5, and 7.6 km, respectively.

antinodes 4 and 5 in Figure 3a. We simply assume the background layer in this interval to be a straight line:

$$n_0(z) = kz + b \quad (z_4 \le z \le z_5) \tag{13}$$

Here, k and b are the parameters for the line; z_4 is the height where antinode 4 is located; z_5 is the height where antinode 5 is located; So formula (11) can be written as:

$$k \cdot \left[\frac{(z+\gamma H)\cdot Ae^{\beta z}\cos\left(\omega t - \vec{k}\cdot\vec{r}\right)}{\gamma - 1} - z \right] + b \cdot \left[\frac{Ae^{\beta z}\cos\left(\omega t - \vec{k}\cdot\vec{r}\right)}{\gamma - 1} - 1 \right] \simeq n_s(z) \tag{14}$$

Here, $Ae^{\beta z}\cos(\omega t - \vec{k}\cdot\vec{r})$ is the wave perturbation, and the wave parameters obtained in step 3 can be used to calculate this value. Formula (13) is just a linear formula for k and b. With the values of n_s at the heights where the two adjacent antinodes are located, k and b can be calculated. In this example, we will use the values of n_s at z_4 and z_5 to calculate k and b, and obtain a line between z_4 and z_5. This process is repeated for consecutive antinodes, resulting in the determination of a series of straight line segments of the background layer.

[42] At the heights where linear wave perturbation vanishes, we keep the previous connection layer. At the Na_s heights, we delete the Na_s peak, and just use a line to connect the background layer. Then we obtain a background layer which consists of many discontiguous lines. We smooth this layer by applying a 7–9 point running mean (about 2 km average).

[43] The solid curves shown in Figure 4 are the new calculated background layers. If we examine Figure 4 carefully, we can see that these new background layers are composed of many discontiguous line segments. In Figure 4c, the peak at 92 km has been deleted in the

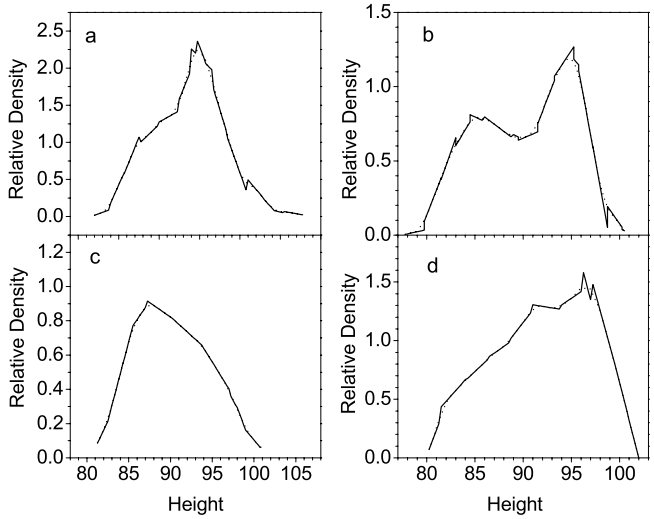

Figure 4. The discontiguous lines (solid curves) calculated from wave parameters obtained in step 3, and the improved background layer (dotted curves) for cases (a) 1, (b) 2, (c) 3, and (d) 4.

Table 4. Correlation Between the Background Layers Obtained in Section 4.4 and Their Corresponding Assumed Background Layer

	Correlation Coefficients for Connection Layer	Correlation Coefficients for the Improved Background Layer	Correlation Coefficients for the Final Obtained Background Layer
Case 1	0.9962	0.9979	0.99998
Case 2	0.9886	0.9941	0.9993
Case 3	0.9943	0.9980	0.99998
Case 4	0.9887	0.9910	0.99976

background layer and a straight line has been used to bridge the gap, as well as that in Figure 4d.

[44] The dotted curves shown in Figure 4 are 2 km running means of the solid curves. From Table 4, we can see they have a higher correlation with the real background layer than the connection layer. So we accept these averaged layers as the improved background layers.

4.4.5. Step 5

[45] In this step we use several Gaussians to fit the improved background layer found in step 4. This is a nonlinear fitting problem to which we can apply the Levenberg-Marquardt method as described below:

[46] The adjustable factors here are background layer parameters, denoted by a_k. From a_k, several Gaussians can be obtained, and an assumed layer are constituted by them. The densities of this assumed layer at each height are given by $y_{i,exp}$, and the densities of the improved layer are given by y_i. The most suitable a_k can be found as follows: $a_k + \delta a_k$, and δa_k can be obtained by solving the linear equations:

$$\sum_i \left(y_i - y_{i,\exp}\right) \frac{\partial y_{i,\exp}}{\partial a_j} = \sum_k M_{j,k}\delta a_k \tag{15}$$

Here,

$$M_{j,k} = \sum_i \frac{\partial y_{i,\exp}}{\partial a_k} \frac{\partial y_{i,\exp}}{\partial a_j}(1 + C \cdot \delta i, j) \tag{16}$$

[47] C is an adjustable constant. We can start the Levenberg-Marquardt scheme with a small value of C. If we overshoot the minimum, the obtained mean square error will increase. Then we back up and increase C. By continuing to increase C, we are guaranteed to decrease the error eventually. If the error decreases, we decrease C again. In this way, the error decreases progressively until a minimum is obtained. A detailed description of Levenberg-Marquardt method is given by *Press et al.* [1988].

[48] The first thing to fit the background layer is to decide how many Gaussians are needed. Here, we try to use the least Gaussians according to the following criterion: the layer obtained from Levenberg-Marquardt method should have the least number of Gaussians, but its correlation coefficient with the improved background layer should be at least 0.999.

[49] Figure 5 shows the Levenberg-Marquardt fit for the background layers in step 4. The solid curves in Figure 5 are the improved background layers in step 4. The dashed curves are the fitted layers, which are composed by the dotted curves.

[50] Shown in Figure 5a is the two Gaussians fit for the improved background layer in Figure 4a, but this fit has some difference at 90 km, and its correlation coefficient is just 0.998. So we add one more Gaussian to the fit and do the Levenberg-Marquardt calculation again. The three Gaussians fit is shown in Figure 5b. We can see this fit is better and its correlation coefficient is 0.9997. So this layer, composed by 3 Gaussians, satisfies our standard.

[51] Shown in Figure 5c is the two Gaussians fit for the improved background layer in Figure 4b. But the correlation coefficient of this fit is only 0.998. Adding one more Gaussian will have better correlation coefficient: 0.9995. So this three Gaussians layer (shown in Figure 5d) can be used to instead of the background layer in Figure 4b.

[52] Shown in Figure 5e is the two Gaussians fit for the improved background layer in Figure 4c. Its correlation coefficient is 0.9996, enough to satisfy the standard.

[53] The background layer in Figure 4d also cannot be substituted by two Gaussians (shown in Figure 5f), as the correlation coefficient is only 0.996. But three Gaussians fit (shown in Figure 5g) has a better correlation coefficient: 0.9990.

4.4.6. Step 6

[54] The "observed layer" in 4.2 is the result of the influence of gravity waves following equation (1), so to extract wave parameters we need to find a reasonable background layer and wave parameters, which have the minimum mean square error between the "observed layer" and that calculated from formula (1). This is a nonlinear fitting problem, which also can be solved by the Levenberg-Marquardt method.

[55] The values of the adjustable parameters found by the Levenberg-Marquardt method depend to some extent on the initial input values. Actually, the obtained wave parameters from the Levenberg-Marquardt method are close to the initial wave parameters (for the wave amplitude, for example, there is generally about 10% difference). So suitable initial parameters are need for the Levenberg-Marquardt fit.

[56] We can use the layer obtained in step 5, constituted by several Gaussians, as the initial input background layer. For the case when Na_s is present, we add one little Gaussian (about the same size of the Na_s peak) at the Na_s height, such as in the last case. The layer in Figure 5g is composed by three Gaussians. Another small Gaussian at 102 km is added, so the initial background layer is composed by four Gaussians. The wave parameters obtained in step 3 can be used as the initial input wave parameters.

[57] After performing the Levenberg-Marquardt fit, the final background layer and the wave parameters are obtained. The final obtained parameters are shown in Table 3.

[58] The final obtained wave parameters for case 1 are: λ_z = 5.29 km, $Ae^{\beta \cdot 92km}$ = 4.92%, and $1/\beta$ = 26.2 km. The final background layer (shown by the dashed curve in Figure 6a) is composed by three Gaussians (shown by the dash-dotted curves in Figure 6a). Shown in Figure 6a is the

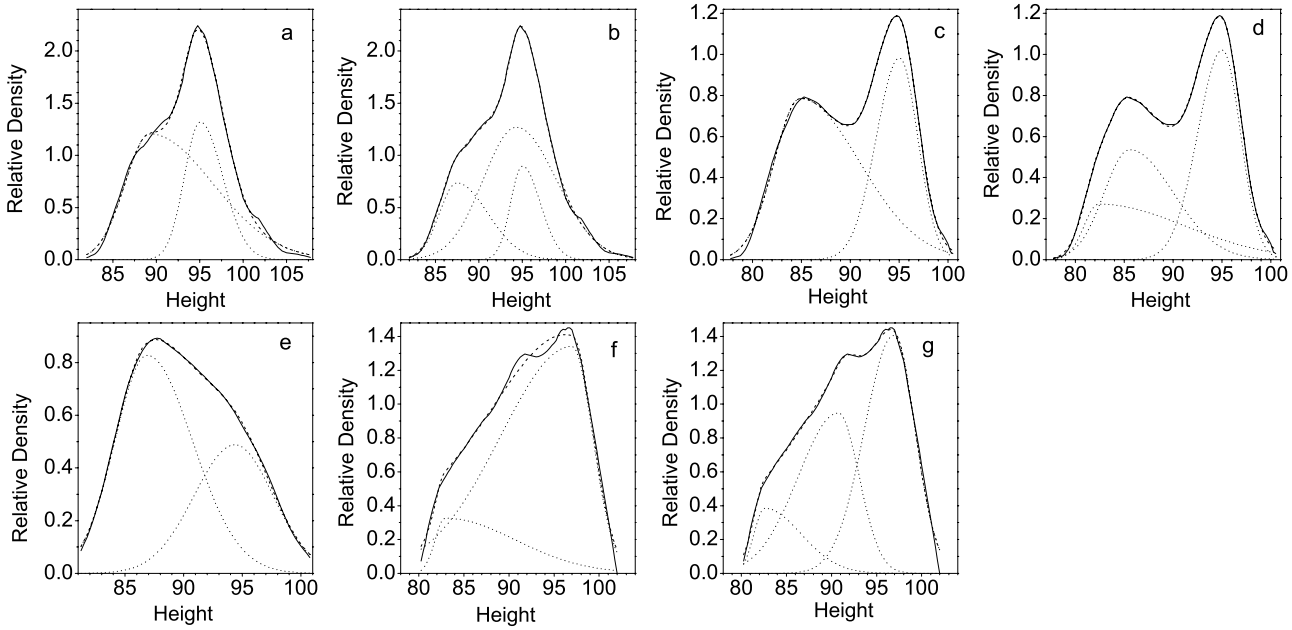

Figure 5. The improved background layer (solid curve) in case 1 is fitted by a layer (dashed curve), which is composed by (a) two Gaussians (dotted curves) and (b) three Gaussians. (c and d) As Figures 5a and 5b but for case 2, (e) as Figure 5a but for case 3, and (f and g) as for Figures 5a and 5b but for case 4.

simulated layer (dotted curve) from the obtained parameters compared with the ''observed layer'' (solid curve), in Figure 6b is the comparison between the sodium density variations from the simulation (dotted curve) and ''observation'' (solid curve), in Figure 6c is the comparison between the wave perturbations, and in Figure 6d is the comparison between the power spectra. We can see from Figures 6a–6d that all the comparisons are very good. Actually, the error of the

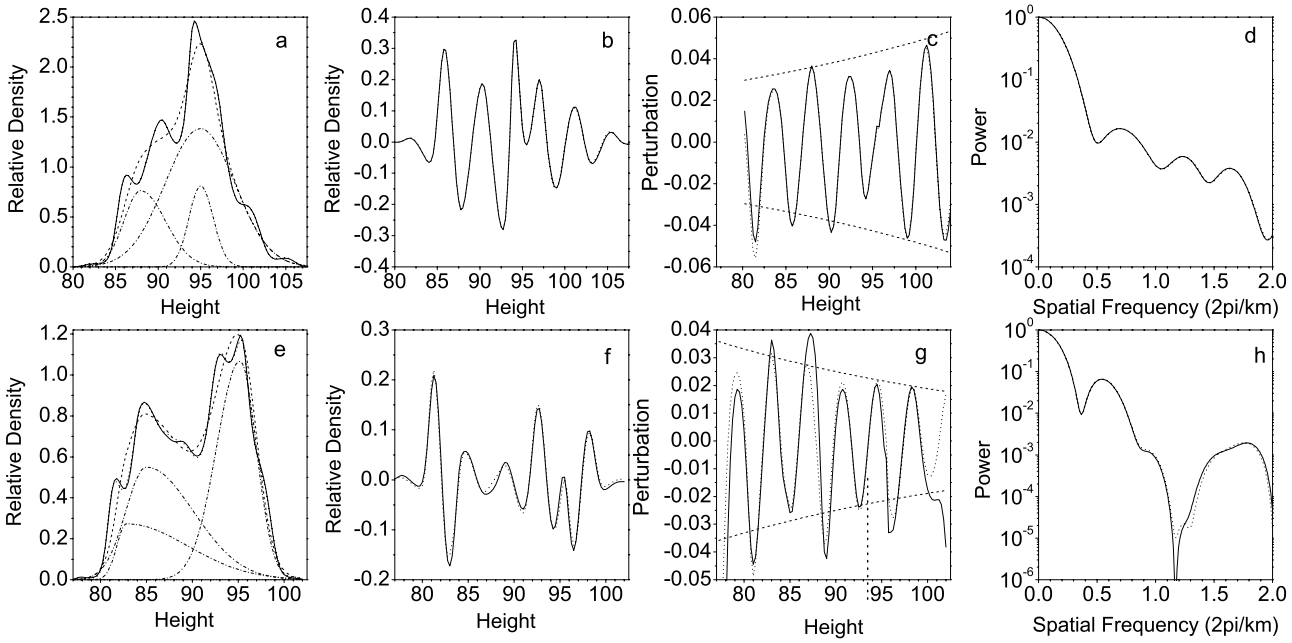

Figure 6. (a) The background layer (dashed curve) composed by three Gaussians (dash-dotted curves) and its corresponding simulated layer (dotted curve) compared with the ''observed layer'' (solid curve) for case 1. (b) The simulated sodium density variation (dotted curve) compared to that from the ''observation'' (solid curve) for case 1. (c) The simulated wave perturbation (dotted curve) compared to that from the ''observation'' (solid curve) for case 1. The dashed curve denotes the wave amplitude variation with altitude. (d) The normalized power spectrum (solid curve) for ''the observed layer'' and the spectrum (dotted curve) for the simulated layer for case 1. (e–h) The same as Figures 6a–6d but for case 2. (The simulation is so close to the ''observation'' that it is difficult to distinguish one from the other.)

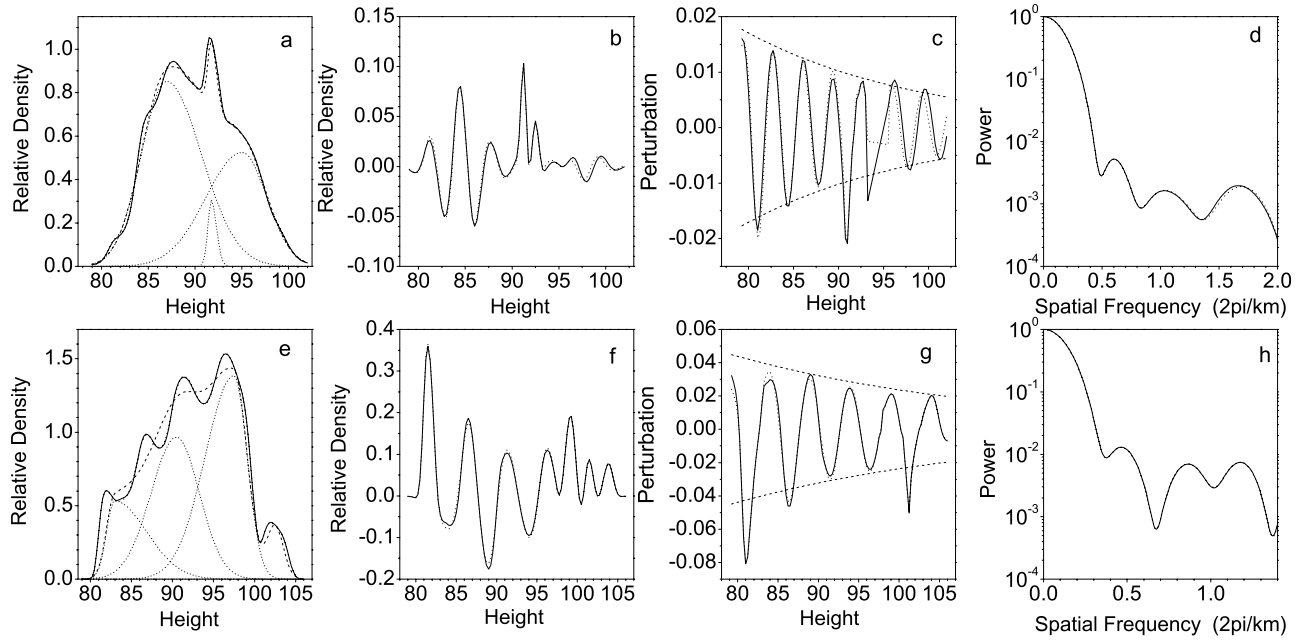

Figure 7. (a–d) The same as Figures 6a–6d but for case 3. (e–h) The same as Figures 6a–6d but for case 4.

wave amplitude is only 1.0%, and the correlation coefficient between the final obtained background layer and the model background layer is as high as 0.99998.

[59] Figures 6e–6h show the comparisons for case 2. Shown in Figure 6e is the simulated layer from the obtained parameters compared with the "observed layer," in Figure 6f is the comparison between the sodium density variations, in Figure 6g is the comparison between the wave perturbations, and in Figure 6h is the comparison between the power spectra. We can see from Figures 6e–6h that all the comparisons are also very good. The difference of the wave amplitude is only 1.7%, and the correlation coefficient is as high as 0.9993.

[60] Figures 7a–7d show the comparisons for case 3. We can see that all the comparisons are good, including the heights where Na_s is presented. The difference of the wave amplitude is 7.7%, and the correlation coefficient is 0.99998. Here, for comparison, the points for calculating the correlation coefficient do not include that at Na_s heights.

[61] Figures 7e–7h show the comparisons for case 4. We can see that all the comparisons are good, including the heights where Na_s is presented. The difference of the wave amplitude is only 1.1%, and the correlation coefficient is 0.99976.

[62] We have described each step of our new method to extract wave parameters. From the above steps, we can see this method mainly considers how to approach the real background layer. From step 1 to step 4 and step 6, the background layer is improved step by step. In case 1, for example, the connection layer has a correlation coefficient (0.9962) with the real background layer, and this correlation coefficient is 0.9979 for the background layer in step 4, while it is 0.99998 for the finial background layer. From Table 4, we can see that the correlation coefficients also increase step by step for other cases. The final obtained wave parameters are also very close to the assumed wave parameters, with an error generally less than 3%.

5. Extraction of Wave Parameters From the Observed Examples

[63] Before extracting wave parameters from the observed layer, we first need to consider how to distinguish a monochromatic wave in the lidar data, as this is the first and basic step when analyzing the data. Here, we adopt the method that the Illinois workers used to distinguish the wave. Only the wavelike structures whose peaks or valleys show the same downward velocity can be considered as monochromatic waves, and this selection criterion eliminates waves suffering serious Doppler shift, thus only waves which experience a uniform Doppler shift or propagate approximately normal to the mean flow are considered [*Beatty et al.*, 1992]. An example which exhibits coherent downward phase progression, as shown in Figure 3 of GV87 or Figure 8a here, gives an idea of how the monochromatic waves are identified from the sodium lidar data.

5.1. Example 1

[64] Shown in Figure 8a is the sodium data sequence observed on 30 May 1996 at São José dos Campos. The time interval between the lidar profiles is 3 min. From Figure 8a, we can see clear downward phase progression structures, which indicate that this is a $\lambda_z \sim 5$ km wave perturbation. We also use the average of profiles 52–54 (solid curve in Figure 8b) to obtain the normalized power spectrum, which is shown in Figure 8h. The power spectrum is seriously distorted. There is a minimum in the power spectrum corresponding to $\lambda_z \sim 5.4$ km. But the maximum having higher spatial frequency is much stronger than that having lower frequency. So this power spectrum does not fit

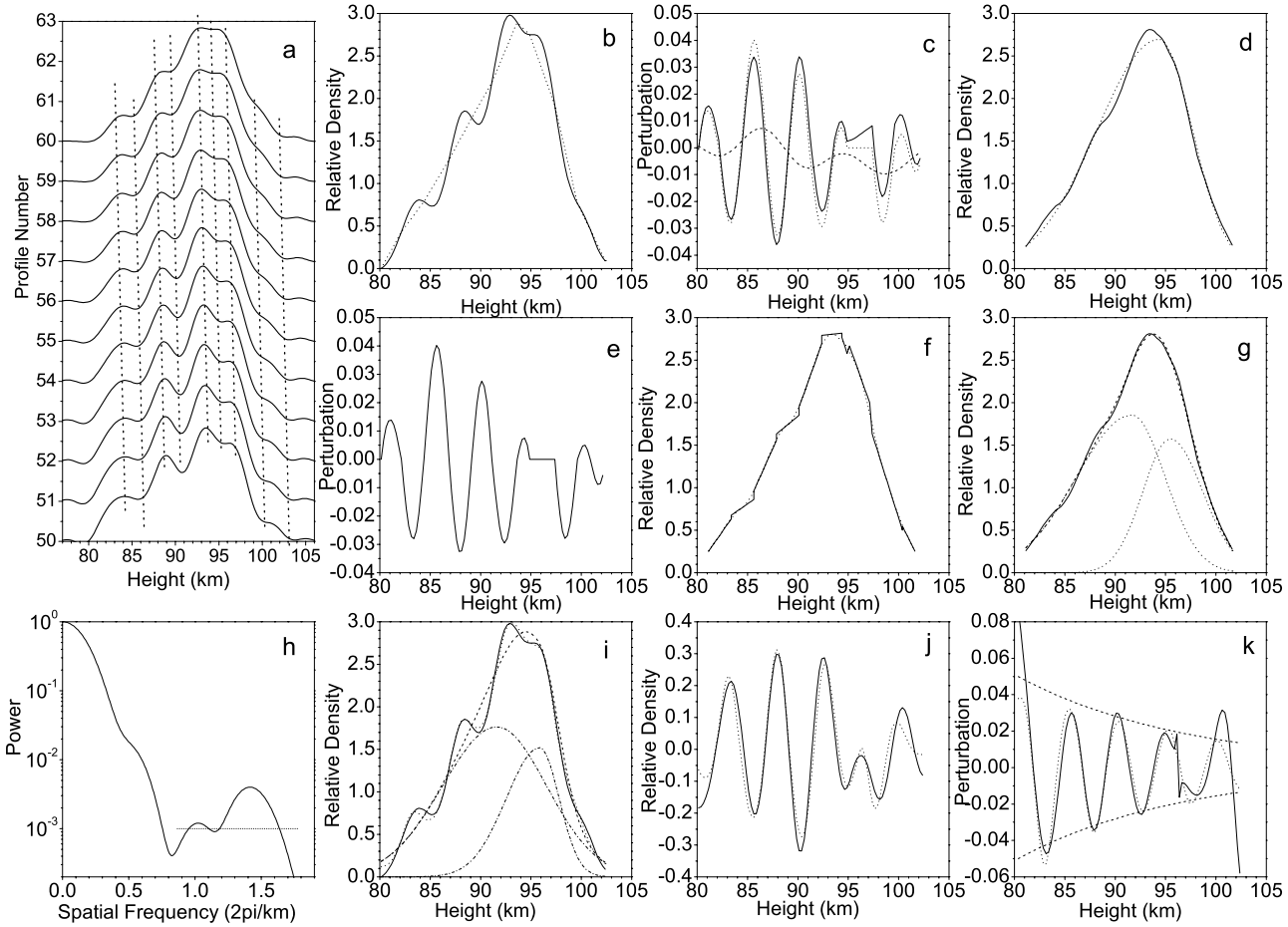

Figure 8. Observation example 1: (a) The sodium lidar data sequence observed on 30 May 1996 at São José dos Campos. The *y* axis is profile number. The time interval between the lidar profiles is 3 min. (b) The average of profiles 52–54 (solid curve) and the connection layer (dotted curve). (c) The wave perturbation (dotted curve) is high-pass filtered with a 6.5 km cutoff to produce a filtered wave perturbation (solid curve) and a discarded low-frequency component (dashed curve). (d) The improved background layer (solid curve) is fitted by a Gaussian layer (dotted curve). (e) The wave perturbation (solid curve) based on the ''observed layer'' and the connection layer. (f) The discontiguous lines (solid lines) calculated from the wave parameters obtained in step 3 and the improved background layer (dotted curves). (g) The improved background layer (solid curve) is fitted by a layer (dashed curve), which is composed by two Gaussians (dotted curve). (h) The normalized power spectrum (solid curve) for the observed layer in Figure 8b. (i) The final obtained background layer (dashed curve) composed by two Gaussians (dash-dotted curves) and its corresponding simulated layer (dotted curve) compared with the observed layer (solid curve). (j) The simulated sodium density variation (dotted curve) compared to that obtained from the observation (solid curve). (k) The simulated wave perturbation (dotted curve) compared to that from the observation (solid curve). The dashed curves denotes the wave amplitude variation with altitude. (The simulated perturbed layer is so close to the observed layer that it is difficult to distinguish one from the other.)

the model illustrated in Figure B1 of GV87, and the GV87 technique cannot be adopted for this case.

[65] Although the GV87 technique cannot be adopted for this case, gravity wave parameters can be extracted by our method:

[66] In step 1, a connection layer can be obtained by connecting the midpoints, as shown by the dotted curve in Figure 8b.

[67] In step 2, from the observed layer and the connection layer, a wave perturbation can be obtained, which is shown in Figure 8e.

[68] From Figure 8e, we can see this wave perturbation roughly corresponds to an overdamped wave. No obvious prominent wave antinode corresponding to the peak in the observed layer is visible.

[69] In step 3, the observed layer indicates a $\lambda_z \sim 5$ km wave perturbation, so we filter the wave perturbation in Figure 8e with a 6.5 km high-pass filter. The filtered wave perturbation is shown by the solid curve in Figure 8c, and it can be used to extract wave parameters.

[70] The average distance between each adjacent antinode of the filtered wave perturbation is 4.85 km, so the wave-

length is 4.85 km. We make a linear fit to all the antinodes of the wave perturbation, and obtain the value of the wave amplitude (2.0%) and the growth factor (−29.0 km).

[71] In step 4, now we create an improved background layer. Under formula (11), with the previously obtained wave parameters and the points on the observed layer corresponding to adjacent wave antinodes, we draw lines between each adjacent antinode, and obtain a layer shape which is composed of many discontiguous lines, shown by the solid curve in Figure 8f. In the range between around 94−98 km, we keep the original connection layer.

[72] We smooth this discontinuous layer by a 2 km running mean. The dotted curve shown in Figure 8f is the averaged layer, which is much smoother. This layer is the improved background layer.

[73] In step 5, then we use several Gaussians to fit the improved background layer found in step 4. We first use one Gaussian to do the fit, and the result is shown in Figure 8d. The fitted Gaussian has much difference to the improved background layer at 90−95 km, and the correlation coefficient is only 0.996. So one Gaussian fit cannot be used to substitute the improved background layer.

[74] Next we use two Gaussians to do the fit, and the result is shown in Figure 8g. The fitted layer is closer to the improved background layer, and the correlation coefficient reaches 0.9991. So we use this two Gaussian fit to substitute the improved background layer.

[75] In step 6, now we can find much better wave parameters and background layer by the Levenberg-Marquardt fit, which is based on formula (1). The two Gaussians fit obtained previously and the wave parameters obtained in step 3 can be used as the initial input parameters.

[76] The final obtained wave parameters for this example are: λ_z = 4.89 km, $Ae^{\beta \cdot 92km}$ = 2.5%, and $1/\beta = -16.8$ km. The final obtained background layer is composed of two Gaussians, shown by the dash-dotted curves in Figure 8i.

[77] Shown in Figure 8i is the simulated layer obtained from the obtained parameters compared with the observed layer. The simulated sodium density variation and wave perturbation are compared with that obtained from the observation, as Figure 8j and Figure 8k shown, respectively. We can see from Figures 8i−8k that all the comparisons are good, so we think we have found both the right wave parameters and the background layer. Such a good agreement, between the simulation and observation, cannot be seen in other published results concerning the gravity wave-perturbed sodium layer which do not consider the true layer shape.

[78] In this example the background layer is composed by two Gaussians, and the RMS width on the top side of the higher Gaussian is only 1.9 km. This steep gradient on the layer top side, together with the long distance between the two Gaussians (about 5 km), distorts the power spectrum in such a way that the GV87 technique cannot be used.

5.2. Example 2

[79] The previous example shows that a background layer with a steep slope on the topside or bottomside can seriously distort the power spectrum, but there exists another kind of background layer that also can seriously affect the power spectrum, as the following example shows. Shown in Figure 9a is the sodium data sequence observed on 4 June 1996 at São José dos Campos. The time interval between the lidar profiles is 3 min. Clear downward phase progression can be seen, indicating that this is a $\lambda_z \sim 5$ km wave perturbation. We use the average of profiles 37−38 (solid curve in Figure 9b) to obtain the normalized power spectrum, which is shown in Figure 9h. The power spectrum is seriously distorted, as although there is a minimum in the power spectrum corresponding to $\lambda_z \sim 4.5$ km, the maximum having higher spatial frequency is much weaker than that having lower frequency.

[80] GV87 technique does not work for this case, but gravity wave parameters can be extracted by the same procedure as example 1. Here we will not repeat the description of the entire procedure, but just indicate the special points for this example:

[81] In step 2, the obtained wave perturbation (shown in Figure 9e) has a prominent antinode at 102 km. But this antinode corresponds to a valley, not a peak, of the observed layer, and it is just an artifice of the connection layer, and not the result of a sporadic layer.

[82] In step 5, we first use one Gaussian to fit the improved background layer, but the correlation coefficient is only 0.987. So we have to use two Gaussians to do the fit again, and the correlation coefficient reaches 0.9995.

[83] The final obtained wave parameters for this example are: λ_z = 4.85 km, $Ae^{\beta \cdot 92km}$ = 2.58%, and $1/\beta = -16.2$ km. The final obtained background layer is composed by two Gaussians, shown by the dashed curves in Figure 9i. We can see from Figures 9i−9k that all the comparisons between observation and simulation are also good.

[84] From Figure 9i, we can see that the smaller Gaussian in the background layer has very steep gradient. Its RMS width at the bottom side is 0.68 km and is 1.33 km at the top side. So the power spectrum is seriously distorted.

5.3. Example 3

[85] We have shown examples of distorted power spectra. However, sometimes the power spectrum is distorted in such a way as to become similar to the spectrum in Figure B1 of GV87. We believe that large errors can ensue if the GV87 technique is applied to analyze this seeming undistorted spectrum. Shown in Figure 10a is a sodium data sequence observed on 1 December 1999 at São José dos Campos. The time interval between the lidar profiles is 3 min. We use the average of profiles 43−44 (solid curve in Figure 10c) to obtain the normalized power spectrum, as the solid curve shown in Figure 10b, which seems similar to that shown in Figure B1 of GV87. Under the GV87 technique, we find a vertical wavelength of 3.68 km, the amplitude (at 92 km) is 2.41%, and the growth height is 8.8 km.

[86] Again we can extract the wave parameter from the procedure described in section 4; here we only indicate the special points for this example:

[87] In step 2, the obtained wave perturbation (shown in Figure 10e) has a prominent antinode near 92 km, which corresponds to the peak of the observed layer in Figure 10b. This antinode is referred to as antinode 8, and produces too large a wave amplitude, resulting in $1/\beta$ being very small, from the average of antinodes 3 and 4 to the average of 7 and 8, about 6.0 km. The distance between antinodes 2 and 4 is 3.5 km, and it is 10.5 km between 4 and 9. So this is a

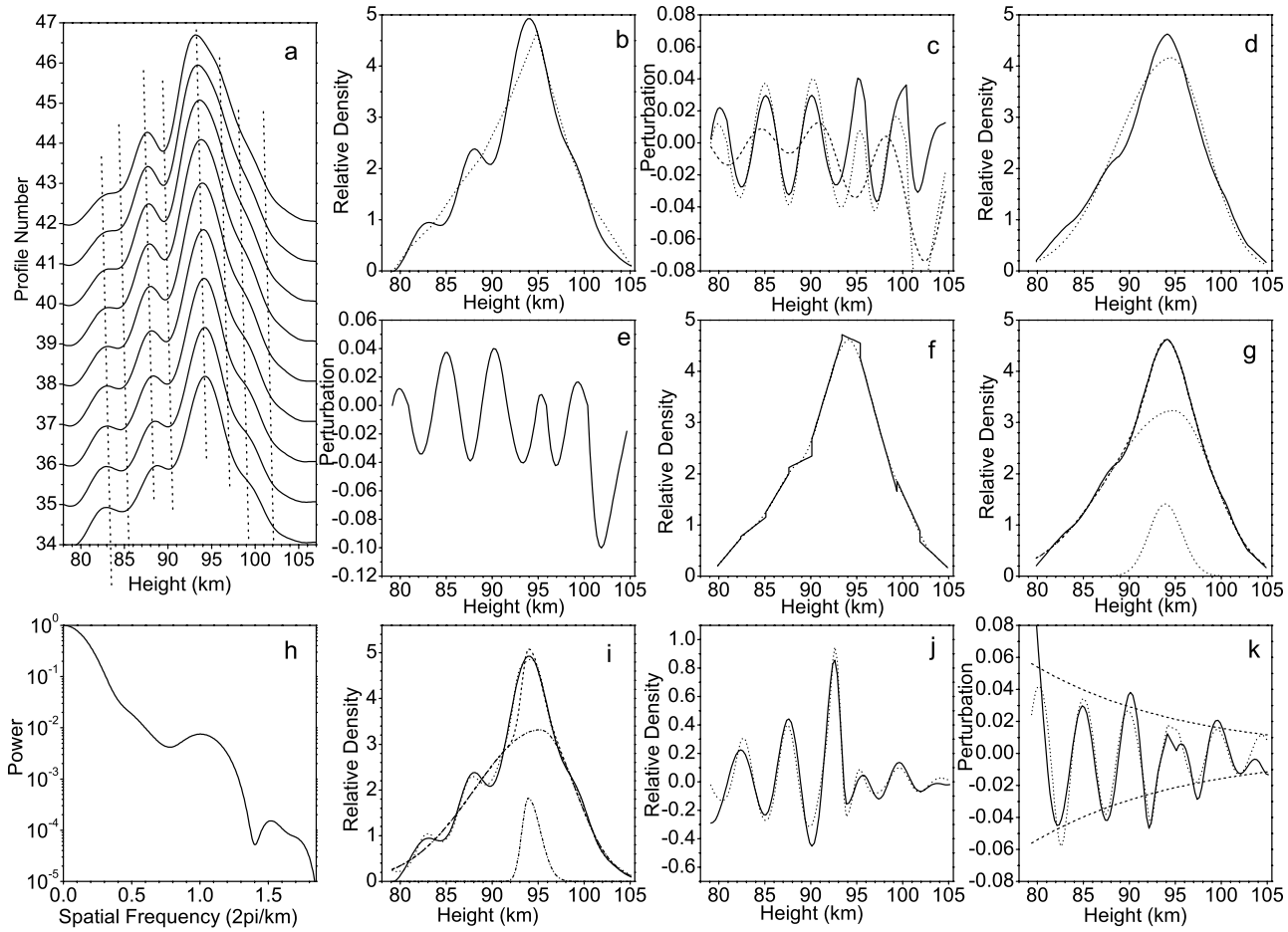

Figure 9. (a–k) The same as Figure 8 but for the observation example 2. The sodium lidar data sequence was observed on 4 June 1996 at São José dos Campos and the time interval between the profiles is 3 min.

$\lambda_z \sim 3.5$ km wave perturbation. But the distance between antinodes 8 and 4 is only 6 km, which is 1 km lower than the expected value. Considering the above reasons, we think the peak in the observed layer near 92 km is just a sporadic layer.

[88] In step 4, at the heights around 92 km, we delete the Na_s peak, and just use a line to connect the background layer.

[89] In step 5, the correlation coefficient of the two Gaussians fit is just 0.998, but three Gaussians fit will produce a higher correlation coefficient: 0.9991.

[90] In step 6, a small Gaussian at 92 km, with a size similar to the Na_s peak, is added to the three Gaussians, and this four Gaussians fit is used as the initial background layer.

[91] The final obtained wave parameters for this example are: $\lambda_z = 3.52$ km, $Ae^{\beta \cdot 92km} = 1.44\%$, and $1/\beta = 1314$ km. The final obtained background layer is composed by four Gaussians, shown by the dashed curves in Figure 10i. Though there is a Na_s peak near 92 km, all the comparisons between observation and simulation are good, as shown by Figures 10i–10k.

[92] In this example, the small Na_s and the broader main layer distort the power spectrum, and change it to be one (the dotted curve in Figure 10h) that seems similar to the spectrum in Figure B1 of GV87. So the wave amplitude extracted by GV87 technique (2.41%) is much higher than that obtained by our method (1.44%).

5.4. Example 4

[93] Here is another example in which the distorted power spectrum seems similar to the spectrum in Figure B1 of GV87. Shown in Figure 11a is a sodium data sequence observed on 17 November 2004 at Wuhan. The time interval between the lidar profiles is 3 min. We use the average of profiles 8–10 (solid curve in Figure 11b) to obtain the normalized power spectrum, as the solid curve shown in Figure 11h, which is close to that shown in Figure B1 of GV87. We find a vertical wavelength of 4.59 km, the amplitude (at 92 km) is 3.92%, and the growth height is found to be 11.75 km by using the GV87 technique.

[94] Different wave parameters are obtained from our method:

[95] In step 2, from Figures 11b and 11e, we can see that there is a peak near 102 km and supposing this is a product of the wave perturbation will result in a very large wave amplitude near this height. The growth height, from the average of the antinodes 6 and 7 to the average of 9 and 10, is 5.3 km, shown in Figure 11e. So this unreasonably large peak must be Na_s.

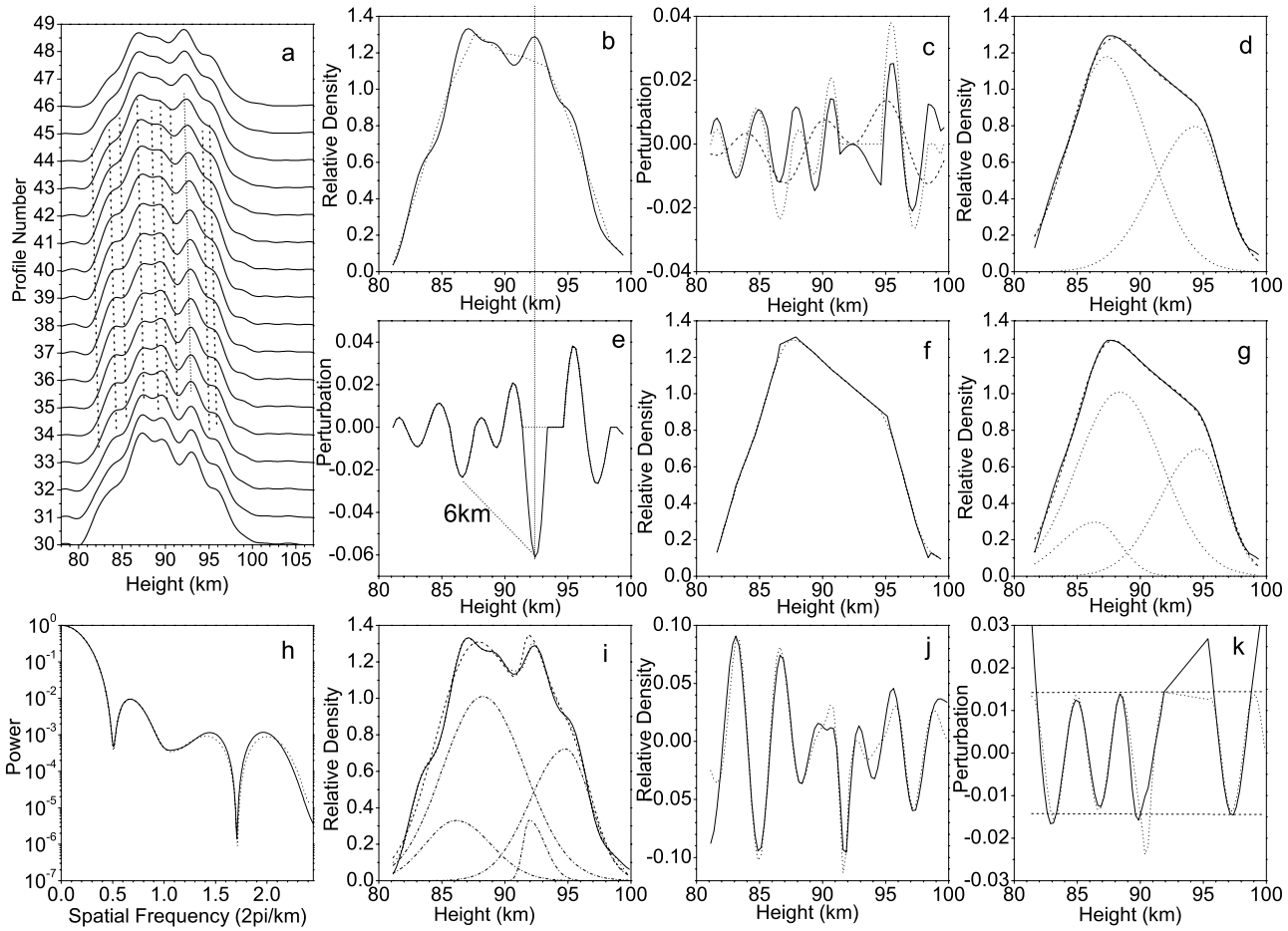

Figure 10. (a–k) The same as Figure 8 (except for Figure 10h) but for the observation example 3. Figure 10h shows the normalized power spectrum (solid curve) for the observed layer in Figure 10b and the spectrum (dotted curve) for the simulated layer in Figure 10i. The sodium lidar data sequence was observed on 1 December 1999 at São José dos Campos, and the time interval between the profiles is 3 min.

[96] In step 4, at the heights around 102 km, we delete the Na_s peak, and just use a line to connect the background layer.

[97] In step 5, the correlation coefficient of the two Gaussians fit is just 0.997, but three Gaussians fit will produce a higher correlation coefficient: 0.9993.

[98] In step 6, a small Gaussian at 102 km, with the similar size to the Na_s peak, is added to the three Gaussians fit, and this four Gaussians layer is used as the initial background layer.

[99] The final obtained wave parameters for this example are: $\lambda_z = 3.67$ km, $Ae^{\beta \cdot 92km} = 1.69\%$, and $1/\beta = -14.37$ km. The final obtained background layer is composed by four Gaussians, shown by the dashed curves in Figure 11i. Though there is a Na_s peak near 102 km, all the comparisons between observation and simulation are also good, as shown by Figures 11i–11k.

[100] In this example, the small Na_s and the steep gradient at the top side distort appreciably the power spectrum, and change it to be one (the dotted curve in Figure 11h) that is close to the spectrum in Figure B1 of GV87. But the wave amplitude extracted by GV87 technique (3.92%) is much higher than that obtained by our method (1.69%).

5.5. Example 5

[101] Examples 3 and 4 show that confusing a sporadic layer with the wave perturbation peak will cause a large error. But occasionally, a very unusual background layer can also distort the power spectrum making it similar to the spectrum in Figure B1 of GV87. Shown in Figure 12a is a sodium data sequence observed on 21 October 1998 at São José dos Campos. The time interval between the lidar profiles is 3 min. We use the average of profiles 23–24 (solid curve in Figure 12b) to obtain the normalized power spectrum, as the solid curve shown in Figure 12h, which is similar to the spectrum shown in Figure B1 of GV87. And we find a vertical wavelength of 4.61 km, the amplitude is 4.2%, and the growth height is 8.8 km by using the GV87 technique.

[102] Much different wave parameters are obtained from the procedure in section 4. Here we just comment on specific characteristics:

[103] In step 5, the correlation coefficient of the two Gaussians fit is just 0.998, but the three Gaussians fit will produce a higher correlation coefficient: 0.9996.

[104] The final obtained wave parameters for this example are: $\lambda_z = 3.91$ km, $Ae^{\beta \cdot 92km} = 1.09\%$, and $1/\beta = -9.76$ km. All the comparisons are also good.

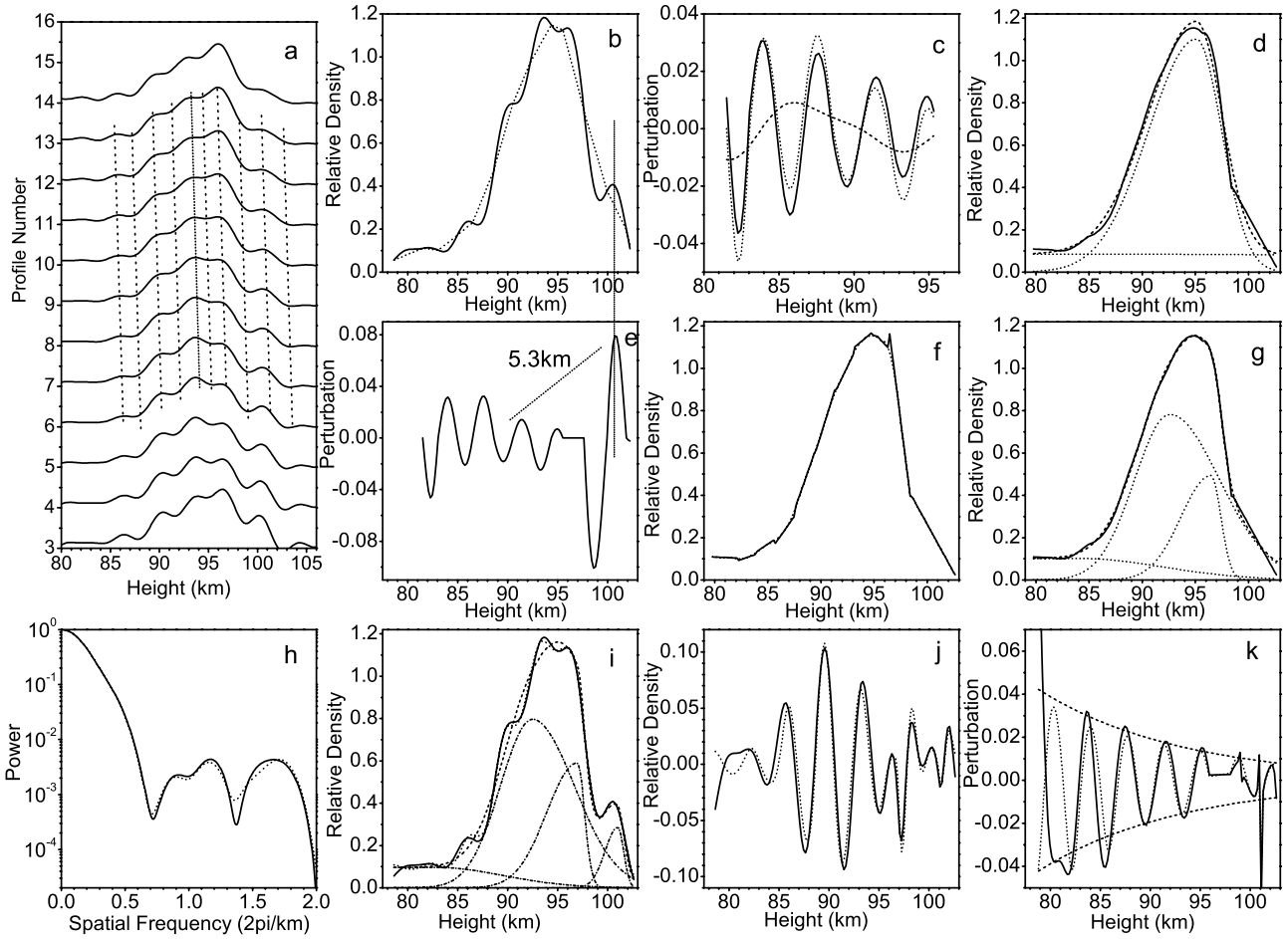

Figure 11. (a–k) The same as Figure 8 (except for Figure 11h) but for the observation example 4. Figure 11h shows the normalized power spectrum (solid curve) for the observed layer in Figure 11b and the spectrum (dotted curve) for the simulated layer in Figure 11i. The sodium lidar data sequence was observed on 17 November 2004 at Wuhan, and the time interval between the profiles is 3 min.

[105] The final background layer obtained comprises three Gaussians, and the distance between the two main Gaussians is about 10 km. The RMS widths of the topside and bottomside are about 1.7 km. This unusual background shape distorted the power spectrum seriously to one which seems similar to the spectrum shown in Figure B1 of GV87 (dotted curve shown in Figure 12h), but applying GV87 technique will obtain a wave amplitude four times larger than that from our method.

5.6. Example 6

[106] The above examples show the seriously distorted power spectrum. However, in our lidar data, there also are some wave perturbation layers whose power spectra are not seriously distorted:

[107] Shown in Figure 13a is the sodium data sequence observed on 17 November 1999 at Wuhan. The time interval between the lidar profiles is 3 min. We can use the average of profiles 97–99 (solid curve in Figure 13b) to obtain the normalized power spectrum, which is shown in Figure 13h. The minimum of the spectrum indicates this is a $\lambda_z = 5.19$ km wave, and in the frequency range near $k_z = 2\pi/5.19$ km, the amplitude of the two maxima is similar. Their averaged spectral value, together with that of the minimum, implies a wave amplitude of 4.96%.

[108] Applying our technique, as described in section 4:

[109] In step 5, the correlation coefficient of the one Gaussians fit is just 0.994, but two Gaussians fit will produce a higher correlation coefficient: 0.9990.

[110] The final obtained wave parameters are: $\lambda z = 5.23$ km, $Ae^{\beta \cdot 92km} = 4.69\%$, and $1/\beta = 32.7$ km. We can see from Figure 13 that all the comparisons are also good.

[111] The background layer in this example is near to a single Gaussian, so the power spectrum is just lightly distorted and the wave parameters extracted from the GV87 technique are similar to those from our method, as Table 5 shows.

6. Statistics of the Distorted Power Spectrum

[112] At São José dos Campos and Wuhan, Na_s occurrence is very high. Also, the background layer is often far from Gaussian. So the power spectrum is often distorted.

[113] We have found 45 monochromatic gravity waves in our 11 years of data, and use our method to extract their wave parameters [*Yang et al.*, 2008]. We also calculated

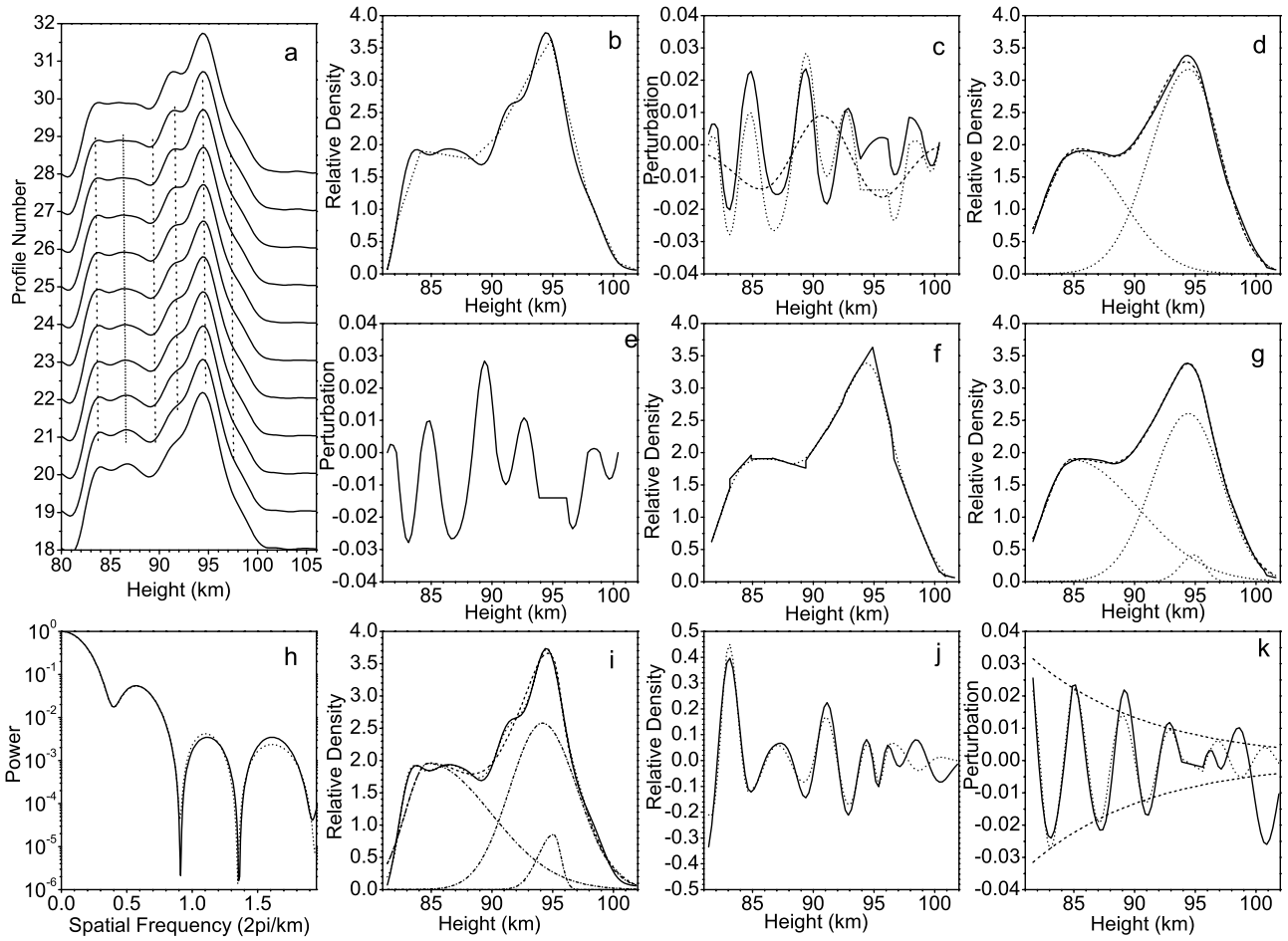

Figure 12. (a–k) The same as Figure 8 (except for Figure 12h) but for the observation example 5. Figure 12h shows the normalized power spectrum (solid curve) for the observed layer in Figure 12b and the spectrum (dotted curve) for the simulated layer in Figure 12i. The sodium lidar data sequence was observed on 21 October 1998 at São José dos Campos, and the time interval between the profiles is 3 min.

their power spectra to see whether they are distorted. Here, we give a summary of our results:

[114] As previously stated, the characteristic spatial power spectrum for a wave perturbed layer is a structure with two maxima and a minimum. We consider the power spectrum to be distorted if the difference between these two maxima is larger than 40%. For example: the spectra in examples 1 and 2 are distorted, while those in examples 3, 4, 5 and 6 not. In the 45 spectra obtained, 13 are not distorted while 32 are distorted, which means only 30% of the wave perturbation layers give undistorted spectra. However, of these 13 cases, in only 8, is the difference between the wave amplitudes calculated from the GV87 technique and from our method less than 25% (such as that in example 6). So only in about 20% of the cases, can the GV87 technique be said to lead to the right wave parameters. In 14 out of a total of 32 cases of distorted spectra a small peak is present in the background layer.

7. Discussion

[115] In step 5, before performing the Levenberg-Marquardt calculation, the initial input parameters, the center heights and RMS widths of the Gaussians, are estimated by personal experience. So this is a "case study." Generally speaking, the nonlinear Levenberg-Marquardt fitting is sensitive to the choice of the initial parameters. But in analysis our 11 years lidar data, we found fitting the background layer by a few Gaussians is not sensitive. We have deigned several initial layers for fit the improved background layers in case 2 and case 3 in section 4. Although these initial layers are much different with the background layer, they produce good fit results. So fitting the background layer by a few Gaussians is not sensitive. For more Gaussians fit to the improved background layer, as the main Gaussians are decided, so the added Gaussians cannot change much.

[116] The connection layer also can be used to design an input background layer. Similar to the above results, we can smooth the connection layer (notice it is not smooth), and use several Gaussians to approach it. The obtained Gaussian layers can be used as the initial layers for step 6. However, the connection is not so closed to the assumed background layer, so this method sometimes will produce much higher errors. We have used this method to extract the wave parameters from the "observed layer" in case 2 in section 4. But the error of the wave amplitude is 5.8%, much higher

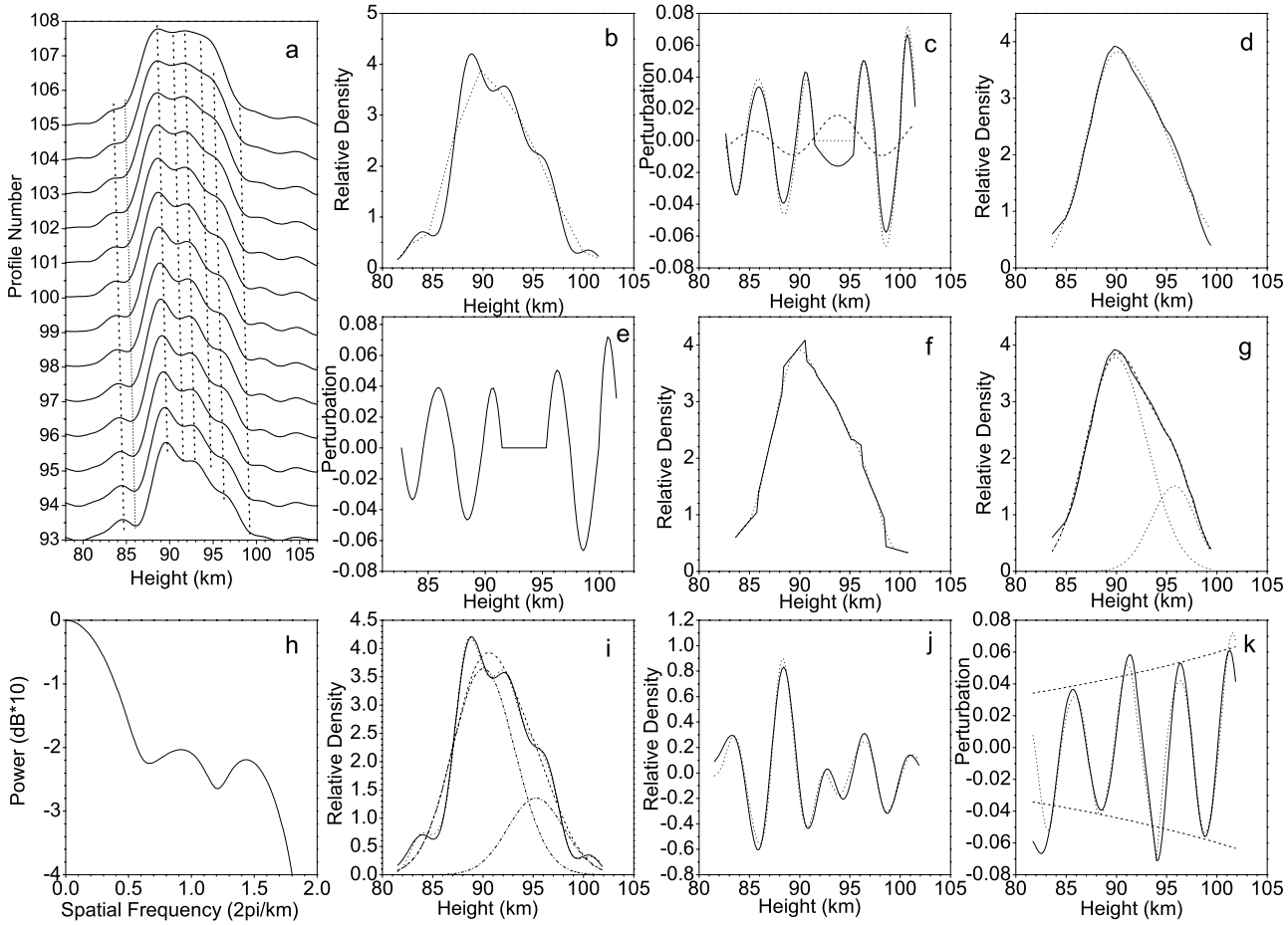

Figure 13. (a–k) The same as Figure 8 but for the observation example 6. The sodium lidar data sequence was observed on 17 November 1999 at São José dos Campos, and the time interval between the profiles is 3 min.

than the error by adopting an improved background layer. So we suggest it is better to adopt step 4.

[117] The typical Na_s, is prominent compared to the main sodium layer, and its duration is often short, so it can be easily recognized. But the minor Na_s, which can last for a long time, are weak compared to the main layer, so they are not easily recognized, as examples 3 and 4 show. Especially the peak in example 4, without the wave analysis, cannot be reliably recognized by a simple visual examination of the sequence of profiles. Also, its power spectrum is close to the spectrum shown in Figure B1 of GV87. At São José dos Campos and Wuhan, this minor Na_s occurrence is very high. So after the connection layer is obtained, the corresponding wave perturbation must be checked to judge whether the peaks in the observed layer relate to wave propagation.

[118] Our method of extracting wave parameters is helpful for distinguishing multiple wave perturbations. Sometimes, the multiple wave perturbation also has an approximate downward phase progression in a few successive profiles. However, the peaks and valleys in this layer cannot be in agreement with those which would be caused by a monochromatic wave, as the wavelengths of the multiple waves are different, and the simulated layer will not be in good agreement with the observed one. Only waves which have very similar wavelengths can produce density variations similar to those caused by a single wave.

[119] As can be seen from formula (1), the interaction of a gravity wave with the sodium layer is quite complicated and highly nonlinear, which makes it difficult to deduce the gravity wave parameters directly from the observed layer. So GV87 uses the spatial power spectrum to extract wave parameters, and the errors caused by the nonlinear term is small, as the contribution of the nonlinear term near k_z in the power spectrum is very small [*Gardner and Voelz*, 1987]. The wave parameters also can be extracted directly from the connection layer and its corresponding wave perturbation (i.e., the wave parameters obtained in step 3). But this step only considers the linear wave perturbation, so errors will ensue. In step 6, the Levenberg-Marquardt calculation is based on formula (1), which includes all the higher terms, so the errors caused by nonlinear terms are eliminated. This is why the final obtained wave parameters are much closer to the assumed wave parameters than the wave parameters obtained in step 3, as shown in section 4.

[120] In our study of the seasonal variation of the gravity wave spectrum [*Yang et al.*, 2006], the background layer was obtained as the average of the observed layer (a whole night in time and 6 km in height), instead of a simulated Gaussian. The background layer could also be obtained by

Table 5. Gravity Wave Parameters Obtained in Section 5

	Obtained in Step 3	Obtained in Step 6	Obtained From GV87 Technique
		Example 1	
λ_z (km)	4.85	4.88	
$Ae^{\beta*92km}$	2.0	2.5	
β (km)	−29.0	−16.8	
		Example 2	
λ_z (km)	4.81	4.85	
$Ae^{\beta*92km}$	3.11	2.58	
β (km)	60.28	−16.2	
		Example 3	
λ_z (km)	3.5	3.52	3.68
$Ae^{\beta*92km}$ (%)	1.67	1.44	2.41
β (km)	16.17	1314	8.8
		Example 4	
λ_z (km)	3.70	3.67	4.59
$Ae^{\beta*92km}$ (%)	1.79	1.69	3.92
β (km)	−13.7	−14.37	11.75
		Example 5	
λ_z (km)	3.90	3.91	4.61
$Ae^{\beta*92km}$ (%)	1.32	1.09	4.2
β (km)	−13.78	−9.76	8.8
		Example 6	
λ_z (km)	5.28	5.23	5.19
$Ae^{\beta*92km}$ (%)	4.50	4.69	4.96
β (km)	25.9	32.7	39.0

using a method similar to that described in this work. But this would not be easy in cases of multiply wave perturbation.

[121] From the variations of the background layer in the above examples, we can know that the background sodium layer actually changes from time to time. This change is probably caused by a sodium source [*Simonich et al.*, 2005]. If the source is layered then Na_s will be present. And if the Na_s is weak and restricted to a limited height range, we can expect to see peaks in the background layer as in examples 3 or 4. These frequently minor peaks in the lidar data indicate the effect of the source. If much fresh sodium is deposited by the source at one part of the layer, the basic layer shape will be changed. And this changed shape can persist for a long time, changing slowly and resulting in an irregular background layer like examples 1, 2 and 5 shown. Long waves, such as tides, can also change the background layer, and this will also affect wave parameter extraction. Indeed, the shape of the layer can change considerably within an interval of ~3–6 h because of the joint effect of diurnal and semidiurnal tides, which means the background layer will change much in this interval, so the wave extraction technique cannot be applicable in this situation. In practice, we only consider the situations in which the averaged sodium layer does not change too much or where the gravity wave period is short.

[122] We assume that the minor peaks are a part of the background layer, but they will produce a large effect on wave parameter extraction when they have a large variation. From simulation, we find that if the variation in the minor peaks is less than 10% or the wave amplitude is very small at the height of the minor peaks, the effect of the minor peak variation can be neglected.

[123] We can use the following method to determine whether or not the temporal variation in the background layer is important. It is the peaks and the valleys in the layer which contribute most in deciding the wave parameters in the Levenberg-Marquardt calculation, and these peaks and valleys correspond to the zero-crossing points 1/4 wave period before. So if the difference in the background layer between the profile used to do the wave analysis and that observed a 1/4 wave period earlier is small, the effect of the background layer variation is not important.

[124] If the wave parameters (vertical wavelength, growth factor) change much with height, the power spectrum will also be distorted. So, quoting GV87 (p. 4675): "Only waves whose spectral signatures fit the model illustrated in Figure B1 have been included in this study. As a consequence, the data are probably biased toward waves are undergoing a uniform Doppler shift throughout the layer or which are propagating normal to the mean flow." In dealing with the variable wave parameters, our method is similar to that of GV87. If λ_z, or β, varies much in height, the peak and valley will change their heights and intensities, and the observed layer cannot agree well with the simulated layer. In our method, these lidar data have also not been included. So, in this respect, our method has the same limitation as the GV87 technique. This method cannot be adopted when the wave parameters vary much with height.

[125] The shortest wave observed by lidar is about 2 km, which is limited by the signal shot noise [*Gardner and Voelz*, 1987]. The longest wave observed by lidar is limited by the sodium layer width. *Gardner and Taylor* [1998] indicated that lidar is incapable of observing the large-scale waves with periods longer than 5 h and both long vertical (>15 km) and horizontal (>1000 km) wavelengths. In the model simulation results [*Xu and Smith*, 2004, Figure 11], we can see a wave perturbation (λ_z = 10 km) cannot be presented so well by the sodium layer. As our method uses density perturbation to extract wave parameters directly, not the power spectrum, it will be affected by the signal noise and the sodium layer width. So we think for our method, the limitation of vertical wavelength is from 2 km to 10 km.

[126] Compared to GV87, this new technique is more widely suitable for extracting the wave parameters, although it is less convenient. Considerable processing time is spent on the multiple calculation steps of the Levenberg-Marquard scheme. The time to perform one Levenberg-Marquard calculation is about 1 min or less.

8. Conclusion

[127] *Gardner and Voelz* [1985, 1987] simplified the analysis of monochromatic waves by assuming the unperturbed layer could be described by a Gaussian function, and proposed a convenient technique for the quantitative estimation of quasi-monochromatic wave parameters. This technique is valid, as it is based upon the fundamental theory of gravity wave perturbation of the sodium layer, and it can be applied to observations with a well behaved background layer, such as at the Illinois site where Na_s is rare.

[128] However, at locations where Na_s occurrence is very high, such as São José dos Campos and Wuhan, the presence of weak Na_s is sometimes not obvious and can be confused with the wave perturbation. Also, the background layer is often far from Gaussian. The irregular shape

of the background sodium layer and the existence of peaks independent of the wave will have a large effect on the derived spatial power spectrum, as our numerical simulations shown. In the GV87 technique, the wave parameters are extracted from the power spectrum, but this technique does not consider the case when the spectrum is distorted. At our site, most spectra (~80%) are distorted, so the GV87 technique cannot be applied. In some cases, although the distorted spectra are similar to the spectrum shown in Figure B1 of GV87, large errors can result from applying the GV87 technique.

[129] To solve this problem, we develop a new analysis method, which starts from a "connection layer." The background layer, as well as the wave parameters, approaches the real layer step by step, as shown by the numerical simulations in section 4. In the observation examples presented in this paper, we show simulated sodium variations and wave perturbations, which agree well with the observations even when the background layer is far from Gaussian. Such a good agreement, between the simulation and observation, cannot be seen in other published results concerning the gravity wave-perturbed sodium layer which do not consider the true layer shape. Our new method is much more widely suitable than GV87 technique, as it can be adopted when the background is often far from Gaussian and the occurrence of minor Na_s is frequent. The limitation of our method is about $2\ km < \lambda_z < 10\ km$.

[130] Our method cannot only be used to extract accurate wave parameters from the lidar data, but also enables us to obtain the true background sodium layer, which may change with time, and often differs greatly from a Gaussian distribution. A study of the unperturbed background sodium distribution should help in understanding the chemical processes which control the layer.

[131] **Acknowledgments.** We gratefully acknowledge Shunsheng Gong for providing some WIPM lidar data and helpful discussion. We also really appreciate the very valuable comments and suggestions of three reviewers. This work has been supported by the Fundação de Apoio a Pesquisa do Estado de São Paulo and the Conselho Nacional de Desenvolvimento Científico e Tecnológico.

References

Beatty, T. J., C. A. Hostetler, and C. S. Gardner (1992), Lidar observations of gravity wave and their spectra near the mesopause and stratopause at Arecibo, *J. Atmos. Sci.*, *49*, 477–496, doi:10.1175/1520-0469(1992)049<0477:LOOGWA>2.0.CO;2.

Chanin, M. L., and A. Hauchecorne (1981), Lidar observation of gravity and tidal waves in the stratosphere and mesosphere, *J. Geophys. Res.*, *86*, 9715–9721, doi:10.1029/JC086iC10p09715.

Clemesha, B. R., D. M. Simonich, and P. P. Batista (1992), A long-time trend in the height of the atmospheric sodium layer: Possible evidence for global change, *Geophys. Res. Lett.*, *19*, 457–460, doi:10.1029/92GL00123.

Clemesha, B. R., P. P. Batista, and D. M. Simonich (1998a), Wave-associated sporadic neutral layers in the upper atmosphere, *Rev. Bras. Geofis.*, *15*, 237–250.

Clemesha, B. R., P. P. Batista, and D. M. Simonich (1998b), Vertical structure in the topside sodium layer, *Geophys. Res. Lett.*, *25*, 3305–3308, doi:10.1029/98GL02563.

Clemesha, B. R., P. P. Batista, and D. M. Simonich (2001), Simultaneous measurements of meteor winds and sporadic sodium layers in the 80–110 km region, *Adv. Space Res.*, *27*, 1679–1684, doi:10.1016/S0273-1177(01)00238-1.

Collins, R. L., A. Nomura, and C. S. Gardner (1994), Gravity waves in the upper mesosphere over Antarctica: Lidar observations at the South Pole and Syowa, *J. Geophys. Res.*, *99*, 5475–5486, doi:10.1029/93JD03276.

Gardner, C. S., and J. D. Shelton (1985), Density response of neutral atmospheric layers to gravity wave perturbations, *J. Geophys. Res.*, *90*, 1745–1754, doi:10.1029/JA090iA02p01745.

Gardner, C. S., and M. J. Taylor (1998), Observational limits for lidar, radar, and airglow imager measurements of gravity wave parameters, *J. Geophys. Res.*, *103*, 6427–6438, doi:10.1029/97JD03378.

Gardner, C. S., and D. G. Voelz (1985), Lidar measurements of gravity wave saturation effects in the sodium layer, *Geophys. Res. Lett.*, *12*, 765–768, doi:10.1029/GL012i011p00765.

Gardner, C. S., and D. G. Voelz (1987), Lidar studies of the nighttime sodium layer over Urbana, Illinois, 2. Gravity waves, *J. Geophys. Res.*, *92*, 4673–4694, doi:10.1029/JA092iA05p04673.

Gibson-Wilde, D. E., I. M. Reid, S. D. Eckermann, and R. A. Vincent (1996), Simulation of lidar measurements of gravity waves in the mesosphere, *J. Geophys. Res.*, *101*, 9509–9522, doi:10.1029/95JD03578.

Kamiyama, H., and F. Tomita (1985), Wave motions in the upper atmospheric sodium layer observed with a lidar technique, *Handbook MAP*, *18*, 253–258.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1988), *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge Univ. Press, New York.

Rowlett, J. R., C. S. Gardner, E. S. Richter, and C. F. Sechrist Jr. (1978), Lidar observations of wave-like structure in the atmospheric sodium layer, *Geophys. Res. Lett.*, *5*, 683–686, doi:10.1029/GL005i008p00683.

Senft, D. C., and C. S. Gardner (1991), Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana, *J. Geophys. Res.*, *96*, 17,229–17,264, doi:10.1029/91JD01662.

Shelton, J. D., C. S. Gardner, and C. F. Sechrist Jr. (1980), Density response of the mesospheric sodium layer to gravity wave perturbations, *Geophys. Res. Lett.*, *7*, 1069–1072, doi:10.1029/GL007i012p01069.

Simonich, D. M., B. R. Clemesha, and V. W. J. H. Kirchhoff (1979), The mesospheric sodium layer at 23°S: Nocturnal and seasonal variations, *J. Geophys. Res.*, *84*, 1543–1550, doi:10.1029/JA084iA04p01543.

Simonich, D. M., B. R. Clemesha, and P. P. Batista (2005), Sporadic sodium layers and the average vertical distribution of atmospheric sodium, *Adv. Space Res.*, *35*, 1976–1980, doi:10.1016/j.asr.2005.06.030.

Xu, J., and A. K. Smith (2004), Studies of gravity wave-induced fluctuations of the sodium layer using linear and nonlinear models, *J. Geophys. Res.*, *109*, D02306, doi:10.1029/2003JD004038.

Yang, G., B. Clemesha, P. Batista, and D. Simonich (2006), Gravity wave parameters and their seasonal variations derived from Na lidar observations at 23°S, *J. Geophys. Res.*, *111*, D21107, doi:10.1029/2005JD006900.

Yang, G., B. Clemesha, P. Batista, and D. Simonich (2008), Lidar study of the characteristics of gravity waves in the mesopause region at a southern low-latitude location, *J. Atmos. Sol. Terr. Phys.*, *70*(7), 991–1011.

P. Batista, B. Clemesha, D. Simonich, and G. Yang, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 12227-010 SP, Brazil. (yyyguotao@hotmail.com)