An investigation of scattering mechanisms and dynamics in PMSE using coherent radar imaging

Tian-You Yu\textsuperscript{a,}\textsuperscript{*}, Robert D. Palmer\textsuperscript{a}, Phillip B. Chilson\textsuperscript{b,}\textsuperscript{2},

\textsuperscript{a}Department of Electrical Engineering and Electro-Optics, University of Nebraska, Lincoln, NE 68583, USA
\textsuperscript{b}MRI Atmospheric Research Programme, Swedish Institute of Space Physics, Kiruna, Sweden

Received 1 September 2000; received in revised form 25 April 2001; accepted 17 July 2001

Abstract

In this work, coherent radar imaging (CRI) was implemented using the ESrange RAdar (ESRAD) system in Kiruna, Sweden (67.88°N, 21.10°E) in the summer of 1998. For the first time, horizontal structures of polar mesosphere summer echoes (PMSE) are studied with high temporal and angular resolution using CRI. Frequency jumps, which represent sudden changes in radial velocity, were observed in several regions during this experiment. Spectrograms and the horizontal structures of PMSE provided by CRI during two cases of frequency jumps, suggest that the frequency jumps were caused by reflecting points originating from a horizontally stratified layer which was modulated by a wave. It was further verified using simulated data and shown that the vertical amplitude of the wave must have exceeded a particular threshold for a given transmit beamwidth. The horizontal wavelength and phase velocity of the proposed wave were estimated to be in the order of 25–34 km and 50–70 m s\(^{-1}\), respectively. Additionally, the vertical amplitude of this wave was up to 400 m, covering several range gates. A hypothesis of phase synchronized multiple layers perturbed by waves is proposed to explain the presence of the frequency jumps at several altitudes simultaneously. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Middle atmosphere dynamics; Remote sensing; Radio science; Interferometry; Signal processing

1. Introduction

The altitude of the mesosphere is too high for balloon-borne observations and too low for in situ satellite measurements. Observations at these altitudes are relatively uncommon and are typically made by rockets, lidar, or high-powered mesosphere–stratosphere–troposphere (MST) radars. Further, the mesosphere contains dynamical processes originating from upper and lower atmospheric coupling. It is also the region where neutral air and plasma both play important roles further complicating the scattering mechanisms (Röttger, 1994a). An example of such complex phenomenon is polar mesosphere summer echoes (PMSE) which are intense radar echoes observed near the polar mesopause during summer months. PMSE are typically 20–30 dB stronger than winter echoes at similar heights and have a peak signal power at approximately 86 km (Ecklund and Balsley, 1981). The scattering/reflection mechanisms used to explain PMSE, which occur over a wide range of radar wavelengths, have been summarized by Röttger (1994b). The mechanisms include Thomson scatter caused by plasma fluctuating at submeter-scales to partial reflection due to steep electron density gradients of a few hundred-meter-scales. More details of PMSE mechanisms...
can be found in the numerous review papers (e.g., Cho and Kelley, 1993; Röttger, 1994b; Cho and Röttger, 1997).

PMSE at VHF wavelengths cannot be explained using traditional turbulent scattering theories since the Bragg scale of the refractive index fluctuations, caused by the electron density fluctuations, is in the viscous subrange at mesospheric altitudes. In other words, turbulence at such small scales will be dissipated by viscosity and therefore, cannot be detected by radars. An exception can occur when the electron diffusivity is reduced due to the presence of ions/aerosols. As a result, the scale of the refractivity index fluctuations can be maintained on a smaller scale and can then be observed by radars of this wavelength (Kelley et al., 1987). The reduction of electron diffusivity could be caused by charged aerosols (Cho et al., 1992) under extremely low temperatures. Indeed, the polar summer mesopause is the coldest region of the atmosphere and the temperature can be further reduced by adiabatic cooling caused by gravity wave modulation (Röttger, 1994a; Cho and Morley, 1995; Rüster, 1995; Chilson et al., 1997).

Another possible mechanism to cause PMSE is the so-called Fresnel reflection or Fresnel scatter, although its origin is still an open question. Fresnel reflection occurs when a single horizontally stratified layer, whose horizontal dimension is much larger than the first Fresnel radius, exists within the radar volume. Normally incident waves are partially reflected by the boundary of the layer. Therefore, it is also called partial reflection. If many of such layers exist within the radar volume, Fresnel scatter is observed (e.g., Röttger and Larsen, 1989). Both mechanisms can be used to explain aspect sensitivity which refers to the dependence of echo power on zenith angle. (e.g., Gage and Green, 1978; Röttger and Liu, 1978). Indeed, aspect sensitivity has been observed within PMSE by several authors (e.g., Czechowski et al., 1988; Cho et al., 1993; Huaman and Balsley, 1998; Chilson et al., 2001b). Although Hocking and Röttger (1997) have shown that partial reflection cannot be caused by the sharp edge of “bite-outs” in the electron density, they did suggest that it may still occur at edges of turbulent layers made up of anisotropic scatterers. Recently, Alcala et al. (2001) have shown evidence that turbulent scattering and partial reflection can occur simultaneously in PMSE within a typical radar range resolution. In their companion paper, Alcala and Kelley (2001) have further shown a good agreement between the observed radar reflectivity and the reflectivity derived from the reflection component of rocket measurements using wavelet analysis.

PMSE have also proven to be an excellent tracer for studies of dynamics in the summer polar mesosphere. One of the most important dynamical processes active near the mesopause is gravity wave motion which has been observed to propagate, steepen, and break in this region. (e.g., Fritts et al., 1988, 1990; Williams et al., 1989; Röttger et al., 1990; Rüster et al., 1996). Upward propagating gravity waves experience a growth in amplitude due to a decrease of the mean neutral density of the atmosphere. Wave amplitude increases until it reaches a critical level where the wave is saturated and the energy contained in the gravity wave is transferred into turbulence to prevent further growth (Fritts et al., 1988; Fritts, 1989). However, in some cases, the energy in the primary wave is transferred into high frequency harmonics that are generated by nonlinear effects. Consequently, the amplitude of the wave can continue growing and the shape of the wave becomes increasingly steepened (Mobbs, 1985; Weinstock, 1986). It has to be emphasized that when wave steepening occurs, turbulence is not necessarily generated. An earlier observation of noctilucent clouds (NLC) made by Witt (1962) shows a case where NLCs are modulated by a monochromatic steepened wave structure. Röttger et al. (1990) have shown sudden frequency jumps in radar spectrograms and suggested that those jumps can be caused by thin scattering layers modulated by steepened waves advecting through the radar beam. Franke et al. (1992) and Chilson et al. (2001a) further supported this hypothesis by using the European incoherent scatter (EISCAT) radar operated in a frequency domain interferometry (FDI) mode to track the motion of such layers.

It has been pointed out by Röttger and Hoz (1990) that the scattering process in PMSE is intermittent and not stationary due to either scatterers in the resolution volume having relative high temporal variability or small distinct structures being advected through the radar beam by the background wind. This ambiguity cannot be resolved using traditional radar techniques. However, coherent radar imaging (CRI) (e.g., Kudeki and Sürúcü, 1991; Woodman, 1997; Palmer et al., 1998), which enhances angular resolution in order to reveal fine-scale atmospheric structures, can possibly provide a solution. The two-dimensional horizontal reflectivity structure within the radar beam can be obtained using CRI. By monitoring the evolution of the reflectivity structures, CRI can be used to determine whether these variations are in the space or time domain.

The goal of this work is to demonstrate the capability and potential of CRI in studying PMSE. A brief review of CRI is presented in Section 2 with the experiment using ESRAD described in Section 3. In Section 4, CRI results from two cases of frequency jumps are presented which suggest the existence of partial reflection or highly anisotropic scatter and wave activity. In Section 5, a numerical simulation is performed and the relationship between the wave shape, the horizontal reflectivity structure, and the frequency jump characteristics is presented and discussed. In addition, a hypothesis of Kelvin–Helmholtz instability (KHI) is proposed to explain the unique features of the observational results. A summary and conclusions are presented in Section 6.

2. An overview of coherent radar imaging

Radar imaging techniques have been widely used in different fields such as radio astronomy, sonar, and seismic
exploration. The use of radar imaging was first introduced into atmospheric applications by Kudeki and Süürüçü (1991) and was termed coherent radar imaging (CRI) by Woodman (1997). Recently, Yu and Palmer (2001) have shown that CRI is a special case of atmospheric radar imaging which was developed to reconstruct the atmospheric structure within the radar volume using multiple receivers and/or multiple transmit frequencies. The visibility function and the angular brightness distribution are fundamental to CRI and form a Fourier transform pair (Thompson, 1986). They can be represented in the following form:

\[ V(l, f) = \int dke^{ik \cdot l} B(k, f), \]

where \( f \) is the Doppler frequency and \( k \) is the wavenumber vector indicating the angular location where the angular brightness is to be estimated and is given by \( k = 2\pi/2[\sin \theta \sin \phi \cos \theta \cos \phi - \cos \theta \sin \phi] \). The zenith and azimuth angles are denoted by \( \theta \) and \( \phi \), respectively, and \( \lambda \) is the radar wavelength. The spatial lag is denoted by \( l \) and is calculated by taking a vector difference between the position vectors of two receivers. The angular brightness distribution, \( B(k, f) \), is the range averaged and normalized signal power density as a function of angle and Doppler frequency (Yu and Palmer, 2001). The visibility function, \( V(l, f) \), is the spatial correlation function and can be estimated by correlating signals from spatially separated receiver pairs. From Eq. (1), the CRI problem can be solved as an inversion problem to estimate the angular brightness given incomplete visibility measurements (Woodman, 1997).

A direct way to obtain a map of the brightness estimate is to take a discrete Fourier transform of the available visibility estimates, which has been termed the Fourier-based method (e.g., Kudeki and Süürüçü, 1991, Palmer et al., 1998). These visibility estimates are spatial samples of the continuous visibility function and have a finite number depending on the number of receivers. Those gaps where no visibility estimates are available are assumed to be zero in the Fourier-based method. Due to such a simplification, this method has inherent resolution limitations. Therefore, other sophisticated methods were introduced to CRI in order to achieve higher resolution and less biased brightness estimates for a given receiver configuration. For example, the maximum entropy (MaxEnt) method takes into account such factors as a priori knowledge of a non-negative brightness and statistical errors in the visibility estimates to produce a map of the brightness which is most likely to be the true brightness distribution as defined in Eq. (1) (Hysell, 1996; Hysell and Woodman, 1997). Another resolution-enhanced method, the Capon method, can be posed as a constrained optimization problem where an optimal weighting function is obtained to suppress interference resulting in superior performance (Palmer et al., 1998). Comparisons of three CRI methods, Fourier-based, MaxEnt, and Capon methods, were made by Yu et al. (2000) and Chau and Woodman (2001) using simulations and observations, respectively.

3. Observational configuration and results

Observations of PMSE using CRI were conducted using the ESRAD VHF system in Kiruna, Sweden (67.88° N, 21.10° E) (Chilson et al., 1999) from July 31 to August 4, 1998. The ESRAD system operates at 52 MHz and has been routinely used to observe PMSE over Sweden (e.g., Kirkwood et al., 1998; Kirkwood and Réchou, 1998). The system can be operated in both the spaced antenna (SA) and the Doppler beam swinging (DBS) modes. The transmitting peak power is 72 kW if the full array 45 m × 45 m is used for transmission. The one-way half-power beamwidth (HPBW) is 7.8°. Six independent, spatially separated receiving arrays are used to sample the backscattered signals. This multiple receiver configuration allows the implementation of CRI algorithms. The antenna configuration and the vertically pointing antenna beam pattern, which would be obtained using Fourier-based CRI, are shown in Figs. 1(a) and (b), respectively. Signals that are near any of the grating lobes will produce aliased structures in the resulting brightness estimates and cannot be distinguished. However, the nearest grating lobe is found at approximately 15° from zenith. Thus, any signal located in a grating lobe would have to be at least 30–40 dB stronger than the signals in the mainlobe to produce spurious structures given the HPBW of the transmitting array of 7.8°, which is unlikely. Nevertheless, the brightness estimates can still be contaminated by sidelobe effects but these effects will be mitigated using Capon CRI. More detailed system parameters of the ESRAD system can be found in Chilson et al. (1999).

The experiment was conducted on August 3, 1998 from 0643 to 2345 UT where strong PMSE returns persisted for more than 10 h. A 1 μs pulse was transmitted every 0.67 ms, with 64 coherent integrations. The transmitting pulse was coded using a 16-bit complementary code. Each data record was stored on magnetic disk after 256 data points for 67 gates and six channels were collected, which corresponds to a dwell time of approximately 50 s. Because the system was operating in the SA mode in order to implement CRI techniques, a three-dimensional wind field could be obtained using full correlation analysis (FCA) (Briggs, 1984). Fig. 2 provides the overall morphology of this PMSE event and the ambient atmospheric conditions over a 10-h period. The echo intensity is represented by the signal-to-noise ratio (SNR), where the noise level is estimated using a polynomial interpolation of the autocorrelation function about zero lag. Only the data for which the SNR is larger than 0 dB are shown and each grid point corresponds to a time and vertical resolution at 50 s and 150 m, respectively. It is possible that this PMSE event may have started earlier than the beginning of the observation. An abrupt change in characteristic occurs at 0940 UT, where both the SNR and vertical
Fig. 1. (a) Antenna configuration of the ESRAD system. The number in each group of four Yagi antennas denotes the receiver to which they are connected. (b) The grating lobe pattern, which is derived based on (a), is plotted from $-30^\circ$ to $30^\circ$ in both the zonal and meridional directions. All grating lobes are at least $15^\circ$ away from the origin.

Fig. 2. Range–time intensity plots of (a) SNR, (b) zonal, (c) meridional, and (d) vertical velocity of a 10-h period on August 3, 1998.
Fig. 3. An example of spectrograms from five range gates and approximately 1 h of elapsed time. Individual spectra are self-normalized at each height and have an aliasing velocity of 16.34 m s$^{-1}$. Diagonal hatching and overlapping structures can be observed in the middle three gates at approximately 1200 UT followed by a disappearance of the signals for a 3-min period. Subsequently, cusp features can be seen at 1215 and 1225 UT.

As mentioned earlier, PMSE are excellent tracers for monitoring dynamics near the summer mesopause. A spectrogram, which is a time history of self-normalized Doppler spectra at a given height, can be used to reveal the radial motion present in the PMSE. Spectrograms have been widely used to study dynamical processes such as wave propagation, steepening, and motions of patches (e.g., Rötger et al., 1990; Cho et al., 1993; Miller et al., 1993; Chilson et al., 2001a). Fig. 3 gives an example of spectrograms from five adjacent heights (87.65–88.25 km) during a 1-h period (1129–1236 UT). The aliasing velocity is approximately 16.34 m s$^{-1}$ with a positive radial velocity indicating upward motion. The colorbar indicates the magnitude of each normalized spectrum given in linear scale. Note that the spectral features generally exhibit sawtooth-like and diagonal hatching shapes instead of sinusoidal oscillations. These were also reported by Rötger et al. (1990) using the
The radial velocity is a superposition of motions. In the case of a wave possessing a relatively short horizontal wavelength, the velocity can be present simultaneously if a wide beam is used or the beam is tilted (van Eyken et al., 1991). Therefore, it is desirable to gain more information about the reflectivity structure within a radar volume to interpret spectrograms. A few attempts have been made to obtain the vertical reflectivity structure using FDI and to incorporate spectrograms to gain more knowledge of the dynamics (e.g., Franke et al., 1992; Chilson et al., 2001a). The present work is the first attempt to obtain the horizontal structure of PMSE at fine-scales using CRI. Efforts will be concentrated on those regions where frequency jumps or overlapping occur in order to utilize the CRI technique to show the horizontal detail so that a better understanding of the mechanisms behind these jumps can be obtained.

A frequency jump is defined by Röttger et al. (1990) as a sudden and noncontinuous change in the Doppler spectrum and is usually accompanied by a change in the sign of Doppler velocity. The first frequency jumps observed in PMSE were reported by Röttger et al. (1988) using the EISCAT radar and were discussed in more detail by Röttger et al. (1990), who proposed two models to explain frequency jumps.

The first model (Model I) consists of a partial reflecting layer modulated by a Gaussian bump or a sinusoidal wave advected through the radar beam. If no background fluid velocity exists, the radial velocity is simply the motion of the reflection point, which is perpendicular to the bore-sight of the transmitting antenna. More than one reflection point can be present simultaneously if a wide beam is used or the wave possesses a relatively short horizontal wavelength. In this case, the radial velocity is a superposition of motions of all reflection points. As a result, diagonal hatching and overlapping features in the spectrograms can occur. In this model, no specific wave shape is required but partial reflection or strongly anisotropic scatter is necessary. This model will be discussed in more detail in Section 5.

In order to lessen the stringent constraint of partial reflection in Model I, a scattering layer was considered in the second model (Model II). A thin, horizontally stratified scattering layer is perturbed by a wave whose horizontal wavelength is much larger than the radar beamwidth. As a result, the radial velocity is the trace of the vertical movement of this thin and wave-modulated layer if no background velocity is assumed. Therefore, the signature of the spectral feature strongly depends on the shape of wave. For a steepened wave structure, it was shown that the radial velocity can jump extremely rapidly from positive to negative when the cusp of the wave passes through the radar (e.g., Franke et al., 1992).

Finally, Cho et al. (1993) proposed a third model (Model III) to explain the overlapping diagonal hatching structures exhibited in the spectrogram using data from CUPRI. If a single localized, reflecting patch advects across the radar beam, a linear radial velocity variation could be obtained in the spectrogram. An overlapping of the frequency can be observed when more than one such patch are present within a radar volume simultaneously.

The diagonal hatching structures exhibited in Fig. 3 from 1150 UT to 1205 UT are now more closely studied. It should be pointed out that overlapping features often observed during frequency jumps are due to the relatively wide beamwidth of the ESRAD system and these features are termed “frequency jumps/overlapping” in this study. In this section, two cases of frequency jumps/overlapping are chosen to demonstrate the capability of CRI and study the relationship between the frequency jumps/overlapping and the horizontal reflectivity structure of the PMSE. Note that frequency jumps/overlapping occurred quite often in this observation but in most cases, CRI was not able to resolve the spatial details because of resolution limitations. Since the purpose of this paper is not to compare the performance of CRI methods, only the results using Capon method are shown; although, the Fourier-based and MaxEnt methods were also implemented. In general, the Fourier-based results do not provide enough resolution and the MaxEnt method produces similar results to Capon method for the cases of interest. However, the Capon method is more computationally efficient than the MaxEnt method.

The first case to be considered is shown in Fig. 4, in which spectrograms, time histories of the peak locations of the brightness estimates in the zonal ($\mu_z$) and meridional ($\mu_\theta$) direction, and maps of the brightness estimates are exhibited from top to bottom, respectively. Spectrograms, $\mu_z$ and $\mu_\theta$ are plotted for a 17-min period from 1154 to 1211 UT at three heights, 87.8, 87.95 and 88.10 km. It is evident that frequency jumps/overlapping are observed at approximately 1201 UT at the three altitudes. The vertical, black,
Fig. 4. Composite plot of results from Case I of frequency jumps/overlapping. The spectrograms are shown in (a) for three adjacent heights (87.8, 87.95, and 88.1 km). The angular location of the peak of individual brightness estimates in the zonal and meridional directions are denoted by $\mu_{x}$ and $\mu_{y}$, respectively, and are shown in (b). Individual values of $\mu_{x}$ and $\mu_{y}$ are denoted by black circles and red triangles, respectively. The missing $\mu_{x}$ and $\mu_{y}$ are those data whose SNR are less than $-5$ dB. The brightness estimates shown in (c) are for the marked region where the frequency jumps are present. Each brightness estimate ($64 \times 64$ pixels) is made from $-7^\circ$ to $7^\circ$ in both zonal and meridional directions every 12 s at each gate using Capon CRI. The results of CRI show that a double-peak structure exists during the frequency jump/overlapping period with single-peak, opposing structures appearing before and after the jumps.

Thick lines denote the region where the corresponding angular brightness are shown in the lower panel. Each brightness estimate was made from $-7^\circ$ to $7^\circ$ both in the zonal and meridional directions using Capon CRI and then normalized to the maximum magnitude of the brightness estimates over the entire time history. The temporal frequency dependence of the brightness distribution as shown in Eq. (1) was eliminated through an average over all frequencies in order to obtain statistically robust images with such a high temporal resolution. In addition, $\mu_{x}$ and $\mu_{y}$ are determined by finding
the angular location of the maximum value in each brightness estimate. The colorbar used for the brightness estimates and spectrograms is the same as for Fig. 3. The three rows of brightness estimates in the lower panel correspond to altitudes of 87.8, 87.95, and 88.18 km from bottom to top, respectively. It should be emphasized that the transmitting beam pattern is not removed from the resulting brightness estimates because spurious peaks will appear at the edge and become difficult to interpret.

In Fig. 4(c), it is interesting to note that the horizontal reflectivity field of PMSE exhibits similar behavior of jumps/overlapping as seen in the spectrograms. Before the frequency jumps/overlapping, the positive radial velocity is associated with a single-peak structure with an apparent motion in the south-west direction (1200:54–1201:18 UT). Subsequently, a second single-peak structure appears from the north-east direction and both structures are present simultaneously exhibiting a double-peak structure during the frequency jumps/overlapping (1201:32–1202:08 UT). After which, only the second single-peak structure remains and moves toward the south-west direction and is associated with a negative radial velocity. It suggests that Model I provides the most appropriate explanation of such transitions exhibited in the brightness distribution and consequently cause the frequency jumps/overlapping for this case. The peak location of each brightness estimate ($\mu_1$ and $\mu_2$) can be thought of as the angular location of the reflecting point which originates from the surface of the perturbed layer. Note that the actual zenith angle of the reflecting point is slightly larger than the peak location because the transmit beam pattern remained in the brightness estimates. Due to the relatively wide beamwidth of ESRAD, two reflecting points can simultaneously exist causing an overlapping spectral feature. It will be shown in the next section that when the reflecting (or highly anisotropic scattering) layer is modulated by a disturbance whose vertical amplitude exceeds a threshold, multiple reflecting points or even the diagonal hatching structures centering at zero can be observed only when small patches continually enter the radar beam from an azimuthal direction of $270^\circ$. In this case, Model III would further imply that there exists a localized source to generate reflecting patches periodically. Otherwise, patches can enter the radar beam from any direction against the wind. Frequency jumps may not occur in such a case, but these small patches can be traced using CRI. Thus, it is unlikely that the structure in the brightness distribution shown in Fig. 4 is caused by reflecting patches as described in Model III. Moreover, Alcala et al. (1995) have shown a case where the frequency jumps were related to the behavior of a wind corner, but our data does not support the existence of a wind corner (Widdel and von Zahn, 1990) in this case.

The second case shown in Fig. 5 is chosen toward the end of this PMSE event during the period 1616–1634 UT at three altitudes, 83, 83.15 and 83.3 km. Frequency jumps/overlapping are evident in the spectrogram at approximately 1623 UT at three altitudes. But the horizontal structure exhibits a broader feature rather than a double-peak structure during the frequency jump/overlapping. Nevertheless, the single-peak structure jumps from the first to the third quadrant before and after the frequency jump/overlapping (most obvious at 83.15 km). It is suggested that a double-peak structure is actually present during the frequency jump/overlapping but the CRI measurements do not have adequate resolution to observed two peaks. Note that the MaxEnt technique was also implemented for this time period but it did not provide the necessary resolution. It is apparent that these single-peak PMSE horizontal structures are closer to zenith and have wider feature than those structures shown in the first case. A possible explanation is that the vertical displacement caused by the wave is not as large as the first case. The relationship between location of reflecting point and wave structure will be discussed in the next section. The wide feature could also be explained by a reflecting layer modulated by a superposition of waves whose shapes are slightly different. Therefore, the layer surface would produce diffuse scattering resulting in a broader structure.

5. Simulation and discussion

In order to further justify Model I, a simulation was employed. It is assumed that a horizontally stratified and reflecting layer exists at a height $H_0$. Further, the layer is perturbed by a monochromatic wave whose horizontal wavelength and frequency are $\lambda_0$ and $\omega$, respectively. As a result, the vertical displacement can be written in the following
Fig. 5. Composite plot of results from Case II of frequency jumps/overlapping. Same as previous figure except for the time from 1615 to 1635 UT and altitudes of 83.0, 83.15, and 83.3 km. In this case, Capon CRI does not seem to have enough resolution to resolve the double-peak structure during the frequency jumps/overlapping. Nevertheless, the locations of the single-peak structures still do jump to the opposite side at the time of the frequency jumps.

form,

\[ z(t) = H_0 + A_z \cos(kx - \omega t), \]

where \( A_z \) is the amplitude of the oscillation and \( k = 2\pi/\lambda_h \) is the horizontal wavenumber. As the wave passes through the radar beam, the reflecting points within the HPBW can be determined numerically. The reflecting point is the location where the ray path is perpendicular to the surface of the modulated layer. In addition, the radial velocity can be estimated by taking a temporal derivative of the total phase path which is twice the distance from the transmitter to the reflecting point (Pfister, 1971). Note that no background velocities (vertical or horizontal) are considered in the simulation.

Vertical displacement, reflecting points \( z \), and radial velocity \( V_r \) resulting from the simulation are shown in
Fig. 6. Simulation results of the reflecting point and radial velocity when a partial reflecting layer is modulated by a deep sinusoidal wave, whose horizontal wavelength is 20 km, advected through a vertically pointing radar beam. The mean altitude of the displacement of the distorted layer is at 88 km with a radar beamwidth of 7.8° corresponding to a horizontal coverage of approximately 12 km. The period of the wave is 10 min and the amplitude of the wave is (a) 80 m, (b) 200 m, and (c) 350 m. When the vertical displacement of modulated layer becomes larger as shown in (b) and (c), multiple reflecting points and spectral features can be observed when the crest of the wave is at zenith. In (c), the diagonal hatching and overlapping structures of $\alpha$ and $V_r$ are exhibited when the wave is significantly steepened.

Fig. 6 for two wave cycles with wave amplitudes of 80, 200, and 350 m. In Fig. 6, the altitude is 88 km and the horizontal wavelength and wave period are 20 km and 10 min, respectively. The radar frequency used in the simulation is 50 MHz. Varying the radar frequency will result in different magnitudes of $V_r$ but the shape will be consistent. The HPBW is selected to be 7.8° to match ESRAD, which corresponds to a horizontal extent of approximately 12 km at an altitude of 88 km. When the wave is shallow ($A_z = 80$ m, Fig. 6(a)), only a single reflecting point and single spectral feature are observed at any given time and no frequency jumps/overlapping occur even though a wide beamwidth is used. The traces of $\alpha$ and $V_r$ exhibit more sinusoidal oscillation as well. Note that the oscillations of $V_r$ and $\alpha$ are out of phase. Further, the oscillations of vertical displacement and $V_r$ have a 90° phase shift. When the amplitude of the wave is significantly increased as shown in Figs. 6(b) and (c), on the contrary, more than one reflecting point can exist simultaneously at a given time, resulting in a frequency jump when the crest of the wave is directly above the radar. Furthermore, it is evident that the variation of $\alpha$ and $V_r$ increase and become linear when the amplitude of the wave increases. Note that the magnitude of $\alpha$ is bounded by half of the HPBW. The diagonal hatching structures of $V_r$ and $\alpha$ are observed when the wave has the largest amplitude as shown in Fig. 6(c). As expected, the simulation model can successfully reproduce the observational results of $V_r$ and $\alpha$. The simulation was also implemented using a Gaussian bump and a simulated steepened wave structure and similar results were observed (Yu, 2000). As a result of these simulations, it can be concluded that diagonal hatching of $V_r$ and $\alpha$ as observed in Figs. 4 and 5 can be produced when the vertical amplitude of the wave is large enough for a given HPBW and partial reflection or strongly anisotropic scatter dominates within the PMSE. Similar work has been done by Meek and Manson (1992) and Röttger et al. (1990). However, no specific HPBW was set to limit the illuminated area of the perturbed layer by Meek and Manson (1992). The finite beamwidth effect was considered in Röttger et al. (1990) where results from a layer modulated by a Gaussian bump were shown.

By assuming the layer is modulated by a monochromatic and sinusoidal wave as shown in Eq. (2), the relationship of the maximum zenith angle of the reflecting point ($\alpha_{\text{max}}$) and wave parameters can be derived. Similar discussions have been provided by Meek and Manson (1992). By finding the maximum deviation of Eq. (2) with respect to $\alpha$, $\alpha_{\text{max}}$ can be determined by the following:

$$\alpha_{\text{max}} = A_z k.$$  

The maximum radial velocity ($V_{r_{\text{max}}}$) observed by the radar can be approximated in terms of wave amplitude and period if a small angle assumption of $\alpha_{\text{max}}$ is made.

$$V_{r_{\text{max}}} = 2V_p \sin \alpha_{\text{max}} \approx 4\pi A_z/T,$$  

where $V_p$ is the phase velocity of the wave.
where \( V_p = \frac{\lambda_b}{T} \) is the phase velocity of the wave and the factor of two is used to correct for the fact that the velocity aloft is one-half of the velocity observed at the ground due to the two-way path effect. Thus, wave parameters such as horizontal wavelength, phase velocity, and amplitude can be determined from Eqs. (3) and (5). For example, in the first case study, \( V_{\text{East}} \) and \( T \) can be determined from spectrograms to be approximately 5 m s\(^{-1}\) and 8 min, respectively. The maximum zenith angle of the reflecting point is approximately 2.8° as calculated from \( \mu_z \) and \( \mu_x \). Consequently, \( A_z, \lambda_h, \) and \( V_p \) are calculated to be approximately 191 m, 24.6 km, and 51.25 m s\(^{-1}\), respectively. These values suggest the scale of the observed wave structure is similar to those evident in Fig. 6(b). In addition, this wave is propagating from the northeast to the southwest because the two reflecting points during a frequency jump should be aligned in the direction of propagation. In the second case, \( A_z, \lambda_h, \) and \( V_p \) can be estimated to be approximately 191 m, 34.38 km, and 71.6 m s\(^{-1}\), respectively, with a propagation direction opposite to that in the first case. Note that the type of the wave could be a vertically propagating gravity wave or a locally generated wave but such a determination cannot be made with the current data. The vertical wavelength of the wave could be derived if the dispersion relation is known such as when an upward propagating gravity wave is assumed and the Brunt–Väisälä frequency is known.

Furthermore, a radar with a narrow beam such as the EISCAT radar can observe only a small portion of the wave described earlier. Therefore, neither double-peak structure nor frequency jumps would be produced by such wave if Model I is applied. On the other hand, a wave with horizontal wavelength of less than a few kilometers is needed to produce a frequency jump in Model I given the beamwidth of the EISCAT radar. It can be shown that the period of the wave is on the order of a few seconds using Eqs. (3) and (5), which is unlikely. Therefore, Model I is not appropriate for explaining the frequency jumps observed at EISCAT (Rötger et al., 1990; Franke et al., 1992) due to its small beamwidth. In addition, it is not clear that these anisotropic scatter or partial reflection needed in Model I can exist at such small scale if the EISCAT radar is used. On the other hand, the ESRAD radar is favorable to observe the frequency jumps caused by the mechanism described in Model I.

In the previous simulation, it was assumed that the wave is confined to a single gate independent of wave amplitude. Therefore, the reflecting point is simply a function of zenith angle. This is not realistic given the previous observational results which showed the vertical variation of the wave (±191 m) can cover approximate three gates if a 1 \( \mu \) pulse is used. If the vertical sampling of radars is taken into account, the reflecting points are not only a function of zenith angle but also a function of altitude. Therefore, only the reflecting points within the same range gate can be used to determine the radial velocity. As a result, frequency jumps/overlapping and multiple reflecting points will exist at higher gates when the crest of the wave is close to zenith. When the crest of the wave is away from zenith, the reflecting points will appear at lower gates and possibly, no reflecting points will occur at higher gates depending on the wave shape. Therefore, frequency jumps/overlapping can only be present at a higher gate for a relatively short period of time. In contrast, observational results presented here show that frequency jumps and double-peak reflectivity structures can appear almost simultaneously at several heights (Figs. 4 and 5). In order to explain these observational results, it is proposed that approximately phase synchronized multiple stratified layers are modulated by large amplitude waves. If multiple layers exist within one range gate, Fresnel scatter is proposed. The simulation was modified to consider the gating procedure and only those reflecting points, which are from the same reflecting layer and range gate, are used to calculate the radial velocity.

The results of the simulation with 15 layers and seven gates are shown in Fig. 7. These layers are randomly distributed within 1 km and perturbed by the same sinusoidal oscillations described in Eq. (2) with parameters similar to those derived from the observations \( (A_z = 200 \text{ m and } \lambda_h = 20 \text{ km}) \). The modulated layers are shown in Fig. 7(a) with apparently thick lines caused by closely spaced layers. The thin, horizontal, solid lines indicate the boundaries of the 150 m gates. Fig. 7(b) and (c) show the variations of \( x \) and \( V_x \) for the middle three gates, which show a marked similarity to the observational results.

One possible mechanism to generate these thin laminar layers is associated with a KHI, which is a dynamical instability in the presence of a hydrostatic stable layer and highly sheared conditions. A typical life cycle of a KH bilow evolves through several stages including an initial single layer, breaking wave, braided structure, stretched filaments, and finally a double layer structure (Browning and Watkins, 1970). The first simulation results of KHI in PMSE were shown and compared to measurements by Hill et al. (1999). In their simulation of KHI, it is found that a large number of fine structures, which are more or less elongated in the horizontal direction, can exist simultaneously within a vertical range of approximate 6 km. The vertical scale is the same as the scale of the billow-like structure in the range-time intensity (RTI) plot of SNR shown in Fig. 2(a) from 1030 to 1500 UT. These elongated fine structures would be produced during the “stretched filaments” stage described by Browning and Watkins (1970), where turbulence generated by breaking waves are stretched horizontally by vertical shears. Bolgiano (1968) suggested that the edges of these turbulent layers can become extremely sharp or anisotropic causing partial reflection. But the existence of such edges in the electron gas needs further justification. Recently, Gibson-Wilde et al. (2000) have shown that the double-layer feature of PMSE can be produced due to strong temperature gradients on both sides of single layer for the case of KHI using a direct numerical simulation. Although we do not have further
Fig. 7. Simulation of multiple layers which are synchronously modulated by a deep sinusoidal wave. As a result, multiple reflecting points and overlapping features in radial velocity occur at several gates. The 15 layers are randomly distributed over approximately 1 km (87.5–88.5 km). The modulating wave has a horizontal wavelength and vertical amplitude of 20 km and 200 m, respectively. Seven range gates with 150 m resolution (thin solid lines) are used to determine in which gate the glint angles exist. The middle and bottom plots are the average glint angle and radial velocity observed in the middle three gates.

observational results to support the existence of the KHI, it seems a reasonable mechanism to explain the presence of multiple layers spanning several range gates.

6. Conclusions

Frequency jumps/overlapping exhibited in spectrograms have been observed by many authors, suggesting the motion of coherent structures across the radar volume and three models have been proposed to explain the observations. For the first time, the instantaneous horizontal structure of PMSE have been shown during frequency jumps/overlapping events using CRI. It is found that a single-peak horizontal structure appears at opposite sides of the sampling volume before and after the frequency jumps and a double-peak structure exists when the spectra split into two features. It is suggested that Model I gives the most appropriate interpretation of the observational results for this case. The horizontal structures in the PMSE are primarily caused by reflecting points from a pre-existing layer which is modulated by a wave with a vertical amplitude of approximately 400 m. The source of the wave is uncertain and the wave shape could be steepened although such a hypothesis would require further justification. Although the cause of reflection within PMSE is still open, the relatively narrow spectra and the small width horizontal structure shown in the two cases presented here support this hypothesis.

In Section 5, simulations were used to study the relationship of the wave shape, reflecting points, and radial velocities during frequency jumps. This model successfully reproduced results similar to the observations. The phase velocity, horizontal wavelength, and vertical amplitude of the wave can be determined if a monochromatic wave is assumed. Furthermore, a scenario of phase synchronized multiple layers modulated by large-amplitude waves was proposed to explain why the frequency jump and the double-peak horizontal structure could be present in several range gates simultaneously. The simulation was modified to allow multiple layers distributed within a defined region and the results show extremely good agreement with the observation. The existence of multiple layers can be verified possibly through the recently developed range imaging technique which allows details in the vertical structure to be revealed within one range gate (Palmer et al., 1999; Luce et al., 2001). However, the source of these multiple layers is still in question. One explanation is that these layers could be generated at the latter stage of a KHI, where relatively well-mixed turbulence is stretched by vertical shears and a complex fine-scale structures can be produced. The reflection may come from anisotropic turbulence at the edges of these turbulent layers. Although no direct evidence shows that PMSE layers are produced by a KHI, observations and simulations have been shown to support this hypothesis (e.g., Pan and Röttger, 1996; Hill et al., 1999; Gibson-Wilde et al., 2000).

In this work, CRI has been successfully implemented for the observation of PMSE to study scattering/reflection mechanisms and wave dynamics. However, it should be emphasized that this study is preliminary in nature and that much more information can be extracted from the brightness distribution. For example, the mean angular location of the brightness can be referred to the mean angle of arrival of backscattered power; the width of the brightness distribution can be related to aspect sensitivity. Chilson et al. (2001b) applied a two-dimensional Gaussian fitting algorithm on each brightness estimate to obtain a set of quantitative measurements of the reflectivity structures and to study aspect sensitivity and atmospheric stability within PMSE. This work
and Chilson et al. (2001b) have demonstrated the feasibility of CRI in PMSE studies.

Acknowledgements

R.D.P. and T.-Y.Y. were supported by the Division of Atmospheric Sciences of the National Science Foundation through Grant ATM 99-08616. P.B.C. was supported by the environment and Space Research Institute (MRI) in Kiruna. ESRAD is jointly financed by the Swedish Natural Resource Council (NFR) and the Swedish Space Corporation.

References


