Variability in Viking ion and electric field observations detected by a wavelet filtering technique

L. Liszka
Swedish Institute of Space Physics, Umeå Division, Sörfors, Sweden

M. Øieroset1 and B. Hultqvist2
International Space Science Institute, Berne, Switzerland

Abstract. We present a wavelet filtering technique combined with principal component analysis which opens up new possibilities to study the variability in satellite particle and electric field observations. With this method, frequency components normally obscured by spacecraft spin and energy sweep can be extracted from the data. The method is tested on Viking ion and electric field observations during an interval of elevated conics near 13,300 km altitude. After wavelet filtering, similar wavelet magnitude patterns are found for both the ion and the electric field observations. This similarity indicates the close relation between electric field fluctuations and upflowing ions. The common features appear first in the electric field and few seconds later in the ion data.

Introduction

Observations in the seventies provided the first evidence that not only the solar wind but also the high-latitude ionosphere is an important source for magnetospheric plasma [Shelley et al., 1976; Sharp et al., 1977]. The energetic (>~10 eV) ion outflow at mid-altitudes in the auroral region has traditionally been classified as beams and conics. Beams have the outflow directed along the geomagnetic field line, while the outflowing ions in conics are directed at some angle away from the magnetic field. Conics are called elevated if they are accelerated along the field line [Klumpar et al., 1984]. The mid-altitude ion outflow is believed to originate as conics, or perpendicular heating events, at lower altitudes and then move upward [see e.g. Yau and André, 1997 and references therein].

The importance of transverse electric field gradients for the upward escape of ionospheric particles was first discussed by Cole [1976] and Lennartson [1980]. Several studies have later reported on the simultaneous observations of ion outflow and low-frequency electric field fluctuations, and proposed the electric field fluctuations to be the cause of the ionospheric ion escape [e.g., Fälthammar et al., 1987; Hultqvist et al., 1989]. Transverse electric field fluctuations below about 2 Hz are expected to be the most important driver for the topside ion escape [Lundin and Hultqvist, 1989]. It is generally accepted that ion conics are produced by the acceleration of ions by electric field fluctuations perpendicular to the magnetic field. However, since the spacecraft is moving, it has not been possible to show from the simultaneous observations of ion conics and low-frequency electric field fluctuations which one precedes the other in time. We note that observations from active rocket experiments have shown that injections of heavy ion beams cause waves, which in turn are shown to cause ion heating (Pollock et al., 1995).

In the study presented here we introduce a method based on a wavelet filtering technique that can extract crucial temporal variations in particle and electric field observations.

1 now at Space Sciences Laboratory, University of California, Berkeley
2 also at Swedish Institute of Space Physics, Kiruna, Sweden
With this method we can extract temporal variations obscured by the instrument energy sweep and the spacecraft spin. The method is based on wavelet filtering and signal decomposition.

We start out the paper with a description of the observations and continue with an extensive section on the filtering and decomposition method used. The last part of the paper is devoted to a discussion of the relationship between ion outflow and electric field fluctuations based on the results from the wavelet filtering technique.

**Observations**

We use ion and electric field observations from the Swedish satellite Viking. Viking was launched in February 1986 in an eccentric polar orbit with apogee at 13,500 km and perigee at 800 km [Hultqvist, 1988]. The ion instrument onboard Viking monitored the energy intervals from 40 eV to 1.2 keV (the PISP2 instrument) and from 1.2 keV to 40 keV (PISP1). These spectrometers performed rapid energy scans in 32 energy levels within a time period corresponding to a pitch angle change between full cycles of about 5.5° (~0.58 sec) [Sandahl et al., 1985].

The dc electric field (V1) and low-frequency wave (V4L) instrument on Viking used two orthogonal pairs of spherical sensors mounted on 40-m wire booms. The low-frequency electric field waveform can be determined by the V1 experiment up to about 10 Hz, and by the V4L experiment up to 428 Hz. The two experiments gave very similar data in the frequency range ~0.2-10 Hz [Lundin et al., 1990], which covers the interval of interest in the present study.

**The Wavelet Filtering Technique**

To illustrate the wavelet filtering technique we apply the method on a sample of data from the Viking ion instrument (PISP2) shown in Figure 1. We have plotted 2048 count rate values from orbit 685 giving approximately 40 seconds of data. Measurements of the variability in particle data are influenced by the experimental conditions, such as the spacecraft spin and the energy sweep. Possibilities to study the temporal variability in the ion data are limited by the spin and energy sweep periods. However, using a wavelet filtering technique temporal variations in the particle flux with frequencies higher than those of the spin and energy sweep can be extracted.
Figure 1.  PISP2 counts from Viking orbit 685, June 26, 1986, starting at 09:51:38:636 UT, covering approximately 40 sec of data.

The 40 seconds of Viking ion data in Figure 1 cover two spacecraft spins. During a spin period the ion detector passes through the entire range of pitch angles from 0° to 360° and measures the ion counts for all 32 energy intervals. It can be seen from the figure that the largest source of variance in the count rate is the energy sweep with a period of 0.58 sec. The second largest source of variance is the spacecraft spin with a period of roughly 20 sec. The real temporal variations in the particle flux are the third largest source of variance. When the particle flux is very low, the Poisson statistics of single particle events may become more important than the temporal variations. In a case of high counting rates, as in Figure 1, the Poisson statistics may be neglected.

**Wavelet transform**

The wavelet transform has become a powerful tool for frequency analysis, in particular for non-stationary time series. Discussions of the wavelet transform and its applications can be found in a number of recent books and review articles [e.g., Chui, 1992; Chui et al., 1994, Farge, 1992]. The wavelet transform of a function y(t) is defined as (* here denotes complex conjugation)

\[ w(a,b) = a^{-1/2} \int_{-\infty}^{+\infty} y(t) g^*((t-b)/a) \, dt \]  

(1.1)

where variable \( a \) is the scale dilation parameter and \( b \) the translation parameter. Both parameters are dimensionless. The real- or complex-valued function \( g(t) \) is called a mother (or analyzing) wavelet. Here a particular wavelet transform, the Morlet wavelet, will be used. The Morlet wavelet, being a locally periodic wavetrain, is related to windowed Fourier analysis. It is obtained by taking a complex sine wave, and by localizing it with a Gaussian (bell-shaped) envelope. The Morlet wavelet is defined as:
\[ g(t) = \exp(i\omega_0 - t^2/2) \quad (1.2) \]

and its Fourier transform as:

\[ G(\omega) = \sqrt{2\pi} \exp[-(\omega - \omega_0)^2/2] \quad (1.3) \]

The Morlet wavelet gives the smallest time-bandwidth product [Lagoutte et al. 1992].

\( \omega_0 \) is a phase constant (in the present study \( \omega_0 = 5 \)). For large \( \omega_0 \) the frequency resolution improves, though at the expense of decreased time resolution. In the present study dilation number 1 corresponds to the highest frequency (a half of sampling rate). The highest dilation number corresponds to the lowest observable frequency.

**Time-Series Decomposition using Wavelet Transform**

Many time series observed in physics consist of a deterministic part with a superimposed stochastic component. A powerful technique to separate both components has been proposed by Farge [1993] and implemented in software by Wernik [1997]. In that method, being a kind of non-linear filtering, a wavelet frequency spectrum of the time series is calculated. The time series is decomposed into a deterministic and a stochastic part. A deterministic “strong” part is obtained by setting to zero all wavelet coefficients less than a certain threshold level. Similarly, a stochastic “weak” part is obtained by setting to zero all wavelet coefficients greater than that threshold level. The inverse wavelet transform is then used to calculate the corresponding time series for both the deterministic and the stochastic part. New wavelet spectra are calculated for each partial time series.

The stochastic part must follow a Gaussian probability distribution function. As a measure of departure from a Gaussian distribution the kurtosis (the fourth moment of the distribution) is used. If the threshold is properly selected, the integral of the kurtosis of the stochastic part over the entire frequency range reaches a minimum.

There may be occasions when the method can be applied for a different purpose. That is the case for counting rates with a dominating Poisson statistics modulated with a weak deterministic component. A very low threshold may be used to separate a weak, deterministic component from a strong Poisson component.
Fig. 2a. The total ampligram for the sample of ion counts in Figure 1.

Fig. 2b. The low-20 ampligram for the sample of ion counts in Figure 1. The same type of plot as in Fig. 2a.
Figure 2c. A different method of plotting the same data as in Fig. 2b. The horizontal axis shows now number of samples. The color scale represents natural logarithm of the component magnitude. Only positive portion of the ampligram is shown for clarity, no smoothing.

**Ampligram**

A generalization of the above technique may be used to separate independent components of temporal variations modulating particle observations as those shown in Figure 1. The experience from studies of oscillations in complex mechanical systems indicate that a given oscillation mode usually occurs with a certain intensity. The intensity ratios between possible modes are usually constant in such a system. That observation may be used to generalize the above non-linear filtering technique.

For a time series of N values (N must be an integer power of 2) the following operations are performed

1. A Morlet wavelet transform is performed with at least 128 dilations. Thus, three N x 128 matrices, $A$, $R$, and $I$, are obtained. The matrix $A$ is a matrix of magnitudes of $w_{ij}$:

   $$A = \left| w_{ij} \right| \quad i=1, \ldots, N \quad j=1, \ldots, 128$$

   (3.1)

   $R$ and $I$ contain respective real and imaginary parts of $w_{ij}$.

2. Instead of using the low-pass or high-pass filtering of wavelet coefficient magnitudes, as described in §2, a kind of band-pass filtering of wavelet coefficient magnitudes is used. The entire range of coefficient magnitudes: 0 to $w_{\max}$, or its lowest 20%, is divided into $M$ intervals such that the $k$-th interval is limited by:

   $$w_{\max} \times \frac{(k-1)}{M} \quad \text{and} \quad w_{\max} \times \frac{k}{M} \quad \text{where} \quad k=1, \ldots, M$$

   (3.2)

   Two sets of intervals are used in the present work: 10 equal intervals between 0 and 100% of $w_{\max}$ and in the other set 20 equal intervals between 0 and 20% of $w_{\max}$. The interval 0 - 20% is especially useful for studying weak signal components. For each $k$ the coefficients outside the range defined by (3.2) are identified and zeroed in matrices $R$ and $I$, creating two new matrices $R_k$ and $I_k$. The inverse wavelet transform is performed using $R_k$ and $I_k$ and a new version of the original time series, $y_k(t_i)$ is created. $y_k(t_i)$ is what the signal would look like if only a narrow range of spectral densities would be present in the signal.
3. The operation is repeated M times at 1 or 10% intervals over the interesting range of coefficient magnitudes, usually 0 - 20% of maximum wavelet coefficient magnitude. A real-valued matrix \( M \), consisting of M columns and N rows is created:

\[
M = \{ y_k(t_i) \}
\]

Each column corresponds to a time series which would be observed if only a narrow range of coefficient magnitudes would contribute to the observed signal. A 3-D plot of the matrix \( M \), called here an ampligram, may be constructed. An ampligram covering 0 - 100% of coefficient magnitudes is called here the total ampligram while the ampligram covering 0 - 20% of coefficient magnitudes is called here the low-20 ampligram. In order to investigate the presence of weak components in the signal the low-20 ampligram is a useful tool. The summation of the matrix \( M \) over \( k \) should result in the original sample \( y(t_i) \), if there would be no energy leakage from outside the filter band (3.2). Fig. 2d shows results of summation (red line) of seconds 700 - 800 of the ampligram in Fig. 2c together with the measured data (blue line). It may be seen that the differences between the observed and reconstructed data are of the order of 10%, which means that the energy leakage from outside the pass-band is not very important.

![Fig. 2d. Results of summation (red line) of 2 seconds of data in the the first 5.5 sec of ampligram in Fig. 2c together with the measured data (blue line).](image)

The low-20 ampligram for the data shown in Fig. 1 is shown in Fig. 2b. Another presentation method is used in Fig. 2c. Only positive portion of the ampligram is plotted for clarity. The color scale shows the natural logarithm of the amplitude \( y_k(t) \). The use of logarithmic z-scale enhances the lowest amplitudes. The ampligram demonstrate that the amplitude and phase of different signal components correspond to different spectral densities. The ampligram is a useful way of presenting physical properties of the signal. The ampligram may also be used to determine whether there is one or more semiregular components in the signal. By adding to the problem the wavelet coefficient magnitude (equivalent to spectral density) as a new dimension, different components of the signal, indistinguishable in conventional analysis, may now be separated. A decomposition technique for the ampligram is discussed in the following section.
Principal Component Analysis (PCA)
A multivariate time series consisting of M variables measured at N equally separated instants forms a matrix $M$. The ampligram may be considered as a multivariate time series ($M=20$, $N=1024$ for the above low-20 ampligram) and the principal component analysis (PCA) may be used to identify the number of independent components in the data.

The results of PCA are:
- The vector of eigenvalues of the matrix (latent roots $\lambda_j$), telling how much of the total variance in the matrix may be explained by the consecutive principal components.
- The matrix of component score coefficients $a$, a transformation matrix between the old system of M variables and the principal components (the new coordinate system).
- The matrix of component scores $S$, with one column for each principal component, being a projection of old M variables upon the new coordinate axis (directions of principal components).

The matrix of component scores $S$ is thus the new multivariate time series in the principal component space. It has been found by one of the authors (L.L.) that, in all cases when low-pass filtering was performed in the principal component space, a considerable improvement of signal-to-noise ratio has been obtained without distorting the signal.

Each column of $S$ is low-pass filtered using a simple filter of moving average type. The result of filtering is matrix $S_f$.

After filtering an inverse transform:

$$M_f = S_f \cdot a^{-1} \quad (4.1)$$

is performed resulting in a new version of the matrix $M$.

It is possible to combine the filtering procedure with a decomposition procedure (Liszka, 1997).

If one wants to know what the variations of the M-component vector would be with only one mechanism (or cause), corresponding to the principal component $l$ active, it is possible to mask with zeros all other columns in $S_f$, except column $l$ and to perform a calculation of a new matrix $M_{lf}$:

$$M_{lf} = S_f \cdot a^{-1} \quad (4.2)$$

The operation may be repeated for each component $l$.

As the principal component transformation preserves the variance, the sum of all latent roots, $\lambda_j$, is equal to the total variance. If the data is standardized, i.e. normalized to standard deviation for each variable, the sum of latent roots is equal to the number of variables. The magnitude of latent roots is usually expressed in percent of the total variance. If the data contains only pure noise, all variables will be uncorrelated, and the total variance will be evenly distributed between all latent roots:

$$\lambda_l \text{ noise (\%)} = \frac{100\%}{M} \quad (4.3)$$

The real data, measured or computer simulated, are never perfectly uncorrelated and the variance will not be evenly distributed between all latent roots.

When all variables are related to the same common factor there will be one latent root (the first one, corresponding to the first principal component) significantly larger than the value indicated by (4.3).
Since the ampligram in Figures 2b and 2c may be considered as a multivariate time series (M = 20 columns, N = 2048 rows) the principal component analysis may be used to identify the number of independent modes in the data.

Figure 3. Component loadings of the principal components PC1 and PC2 for the ampligram in Figures 2b and 2c.

**Decomposition into Independent Components**

The principal component analysis of the ampligram matrix was performed to identify the number of significant independent components in the data. As a result the matrix of component loadings, being the correlation coefficients between significant principal components of the ampligram and the M columns of the ampligram, is obtained. An example of component loadings is shown in Figure 3. The diagram shows which ranges of amplitude contribute to the two significant components present in the ampligram of Figures 2b and 2c. The component 1 (dominating) corresponds to coefficient amplitudes 7-11% and the component 2 corresponds to low amplitudes of 2-6%.

The non-linear filtering is now repeated once for each observed component, and the bandpass of coefficient magnitude is selected from Figure 3. The result, after the inverse wavelet transform, shows time series corresponding to the significant components. A wavelet spectrum may be calculated for each component. An example of decomposed components and their wavelet spectra for the data in Figures 1 and 2 is shown in Figure 4. The decomposed components are still contaminated by the energy sweep. It is possible to improve the above decomposition procedure by adding to it a band-stop wavelet filtering around the energy sweep frequency.

The ampligram may be used for calculation of average wavelet spectra, one for each coefficient value. The procedure will generate a 3-D graph showing the frequency (or period) on the x-axis, the wavelet coefficient amplitude in percent on the y-axis and power spectral density as the color scale. A graph of that kind shows the average properties of different modes, if such exist, during the sample period, in the present case during 39 seconds. A sequence of four such graphs for orbit 685 is shown in Figure 5. For the selected interval Viking was in the cusp region near 14 MLT. The dominating features observed in the ion data during that interval are elevated ion conics.
Figure 4. Decomposed modes corresponding to PC1 and PC2 together with their wavelet spectra; logarithmic frequency scale.
Corresponding invariant latitude and magnetic local time are given below each graph. Graph b corresponds to the 40 sec of data shown in Figures 1 - 4.

Figure 5. A sequence of consecutive 3-D wavelet time scale spectra of ampligrams showing development of oscillation modes in the particle flux within the cusp region. Corresponding invariant latitude and magnetic time intervals are indicated below each graph. Time increment between graphs: 39 seconds.

**Comparison of Ion Outflow and Electric field**

The observed modulation of the particle flux is likely to be related to the electric field along the particle trajectory [Lundin and Hultqvist, 1989]. The analysis technique described above was applied to Viking ion and electric field data for the same time interval to study the relation between the two. Figure 6 shows an example of 40 sec of simultaneously observed electrons and ions (40 eV - 40 keV) and electric field (E_{1bsp}). E_{1bsp} has the main component directed along the magnetic field line, but a large perpendicular component will also contribute significantly. The electric field data were sampled at the same rate as the particle data (0.019 sec). The ion spectrogram show two elevated conics. The wavelet filtering technique described above was applied to both the ion data and the electric field data and the resulting ampligrams are shown in Figure 7.
Figure 6. Viking electron and ion spectrogram for 40 sec corresponding to two satellite spins (upper panels) and electric field data ($E_{\text{1bsp}}$) for the same interval (lower panel). The electric field plot is from Hultqvist [1991]. Two elevated ion conics are seen.
Figure 7. Ampligrams (lowest 20% of wavelet coefficients) for PISP1 particle counts (upper graph) and for the electric field (lower graph) for the two spins from Viking orbit 186 shown in Fig. 6. The horizontal scale is expressed in sample numbers: 2048 corresponds to 38.9 seconds.

Comparing the graphs in Figure 7 it must be concluded that there is a striking similarity between some dominating features observed in the ion data and in the electric field data.
Furthermore, there are time shifts of the order of a few seconds between the features observed in both graphs. Since Viking is moving essentially perpendicular to the geomagnetic field during the event, the structures seen in the electric field ampligram is probably observed on another field line than the corresponding features seen for the particles, if horizontal drifts are not significant. We can, however, assume that similar electric field fluctuations occur on neighbor field lines. The ion observations show a ‘stable’ region of elevated conics near the 40 seconds of interest. If the particle flux is accelerated by waves observed in the electric field data, the time shift may represent the time it takes to accelerate the ions by wave-particle interaction. It may possibly also relate the modulation of the particles at a lower altitude by the waves, assuming that the ampligram observed by the satellite being representative for the waves in the region of maximum modulation. By processing the ampligrams in Figure 7 it is possible to construct the average distribution of the wave periods as observed in the particle flux and in the electric field data. The result is shown in Figure 8. The graphs indicate, for ions (left graph) a presence of regular oscillations with a period of approximately 5.8 seconds and the wavelet coefficient magnitude of 11% of maximum value. There seems be also another, weaker, component with larger coefficient magnitude and large time scale variations (5 - 10 seconds). For electric field variability (right graph) the same oscillation period is present. However, that component shows considerable magnitude variability (7 - 12%) but a rather stable time scale. It is possible that also here we see two components, but with smaller separation in coefficient magnitudes.

A choice of 20% of the maximum wavelet coefficients is by no means arbitrary. As mentioned above, the major part of the variance in the data is due to the energy sweep and to the spacecraft spin. It was estimated that the real time - space variations of the particle counts represents less than 20% of the variance. We have therefore assumed that the same relation is valid for the wavelet coefficients. The 20% level in the analysis removes most of the uninteresting variance due to the energy sweep and to the spacecraft spin and keeps the real time - space variations.

Acknowledgments. Viking electric field data were kindly supplied by P. A. Lindqvist, Royal Institute of Technology, Stockholm, Sweden. We thank M. André, University of Umeå, Sweden for valuable discussions. The Swedish satellite Viking was financed by grants from
the Swedish Board for Space Activities. The Viking project was managed by the Swedish Space Corporation. The main contractor was Saab Space AB.

The software used in this work has been developed by Pär-Ola Nilsson, Jan Karlsson and Fredrik Rutqvist of the Umeå Division of the Swedish Institute of Space Physics.

References
Lennartson, W., On the consequences of the interaction between the auroral plasma and the geomagnetic field, Planet. Space Sci., 28, 135, 1980.


Yau, A., and M. André, Sources of ion outflow in the high latitude ionosphere, Space Science Rev., 80, 1, 1997.