



RESEARCH ARTICLE

10.1002/2015RS005652

Key Points:

- Algorithms are developed to identify quiet day curves for subionospheric VLF
- Presmoothed FFT algorithm performs best on synthetic data for quiet day curve
- Application to real data successfully identifies space weather events

Correspondence to:

K. Cresswell-Moorcock,
creal158@student.otago.ac.nz

Citation:

Cresswell-Moorcock, K., C. J. Rodger, M. A. Clilverd, and D. K. Milling (2015), Techniques to determine the quiet day curve for a long period of subionospheric VLF observations, *Radio Sci.*, 50, 453–468, doi:10.1002/2015RS005652.

Received 14 JAN 2015

Accepted 9 MAR 2015

Accepted article online 13 MAR 2015

Published online 28 MAY 2015

Techniques to determine the quiet day curve for a long period of subionospheric VLF observations

Kathy Cresswell-Moorcock¹, Craig J. Rodger¹, Mark A. Clilverd², and David K. Milling³

¹ Department of Physics, University of Otago, Dunedin, New Zealand, ² British Antarctic Survey (NERC), Cambridge, UK, ³ Department of Physics, University of Alberta, Edmonton, Alberta, Canada

Abstract Very low frequency (VLF) transmissions propagating between the conducting Earth's surface and lower edge of the ionosphere have been used for decades to study the effect of space weather events on the upper atmosphere. The VLF response to these events can only be quantified by comparison of the observed signal to the estimated quiet time or undisturbed signal levels, known as the quiet day curve (QDC). A common QDC calculation approach for periods of investigation of up to several weeks is to use observations made on quiet days close to the days of interest. This approach is invalid when conditions are not quiet around the days of interest. Longer-term QDCs have also been created from specifically identified quiet days within the period and knowledge of propagation characteristics. This approach is time consuming and can be subjective. We present three algorithmic techniques, which are based on either (1) a mean of previous days' observations, (2) principal component analysis, or (3) the fast Fourier transform (FFT), to calculate the QDC for a long-period VLF data set without identification of specific quiet days as a basis. We demonstrate the effectiveness of the techniques at identifying the true QDCs of synthetic data sets created to mimic patterns seen in actual VLF data including responses to space weather events. We find that the most successful technique is to use a smoothing method, developed within the study, on the data set and then use the developed FFT algorithm. This technique is then applied to multiyear data sets of actual VLF observations.

1. Introduction

Man-made VLF transmissions propagate for long distances with low attenuation in the Earth-ionosphere waveguide between the conductive Earth surface and the lower edge of the ionosphere (*D* region). This manner of propagation is termed subionospheric. The radiation propagates in a modal fashion within the waveguide, along the great circle path between transmitter and receiver, with the received amplitude and phase of the transmissions being a superposition of these modes [Wait, 1996a; Lynne, 2010]. Variations in the waveguide over time change the modal mix, causing variations in the observed signal amplitude at a receiver. The primary, temporally varying parameter of the waveguide is the *D* region reflection height, which varies on a regular diurnal basis with the presence of solar radiation as well as irregularly in response to space weather events. For an overview of historical VLF science see Barr *et al.* [2000].

Diurnal variations in VLF observations have been used to determine the relationship of solar zenith angle to *D* region parameters for the daytime ionosphere [Thomson, 1993] and determine similar parameters for the nighttime ionosphere [Thomson *et al.*, 2007], when solar radiation has a less dominant influence. The most dramatic modal variations occur as the day-night terminator passes across the transmitter-receiver path, with the varying modal superposition causing very deep minima in signal amplitudes. Clilverd *et al.* [1999] related the timing of a set of twilight modal minima to the moving location of the terminator along a path at a height of 75 km. Lightning sferics within the VLF frequency band have also been used to determine daytime and nighttime *D* region parameters [e.g., Han and Cummer, 2010a, 2010b; Shao *et al.*, 2013].

Space weather events, e.g., solar flares, energetic electron precipitation (EEP) from the radiation belts, and solar proton events (SPEs), can significantly increase the ionization rate in the upper atmosphere, which leads to a reduction in the *D* region height and thus a perturbation in the received VLF amplitude [Clilverd *et al.*, 2009]. To study the effect of space weather events on the ionosphere using VLF propagation, it is important to have a method of determining the undisturbed diurnal variation in the VLF observations, known as the quiet

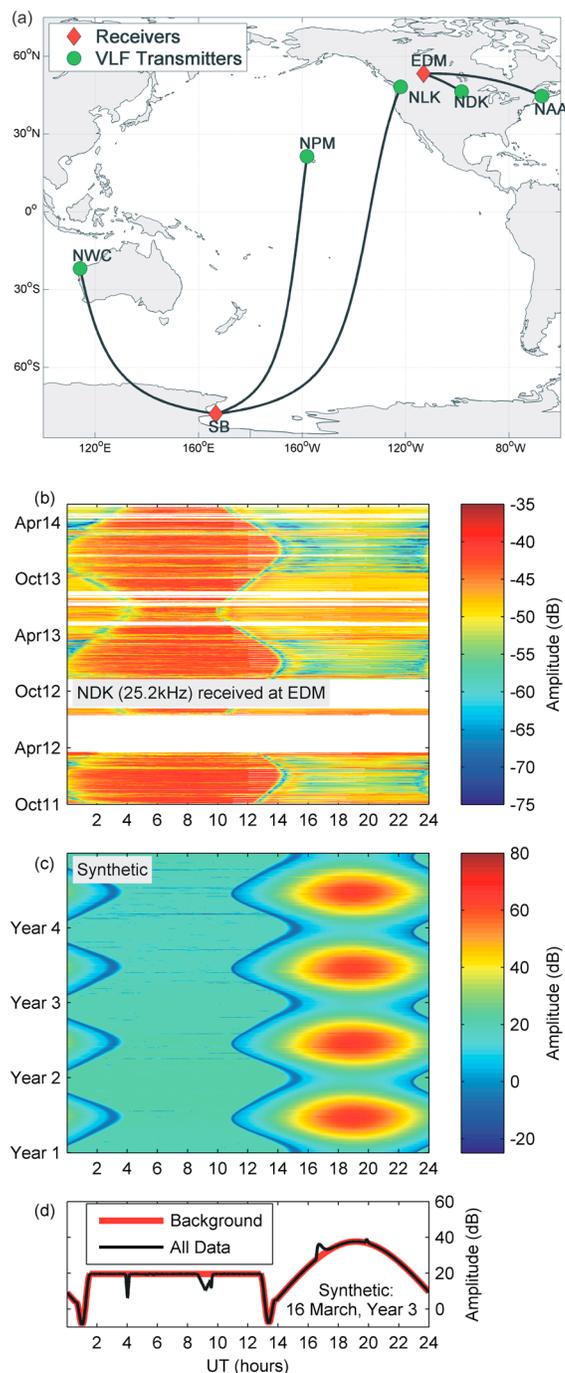


Figure 1. (a) The great circle paths of the AARDDVARK observations analyzed in this study. Green circles indicate the locations of the monitored VLF communications transmitters, with call signs indicated. Red diamonds indicate the locations of the two AARDDVARK receivers (SB: Scott Base, Antarctica, and EDM: near Edmonton, Canada). (b) Observations of the NDK transmission received at the EDM antenna from October 2011 to May 2014. (c) A synthetic data set used for analysis of the success of our QDC finding techniques. White areas in Figure 1b show the placeholder values replacing unusable data. The color scales for the two middle plots are shown to the right of each plot. (d) Data from a representative day of the synthetic data set. The red line is the true QDC, or background, and the black line is the complete data, combining perturbations and background.

day curve (QDC). Once the QDC is determined, the perturbations caused to the VLF signals by these events can be quantified.

In this paper we report on the development of three algorithmic QDC finding techniques for long-period subionospheric VLF data sets. The first technique is based on the QDC technique from *Simon Wedlund et al.* [2014], who created their QDC from the combined curve of several quiet days prior to a period of geomagnetic disturbance that they were investigating. The second technique follows the QDC finding technique based on principal component analysis reported by *Collier* [2009] and *Wautelet and Warnant* [2012]. These studies selected the principal components accounting for the most variance in their data sets for transforming back to data space. The third technique uses a two-dimensional fast Fourier transform to identify the discrete spectrum of the data set before applying restrictions developed within this study to the spectrum and inverting the transform to provide the QDC. These techniques all result in QDCs that reproduce both the diurnal and annual amplitude variations in the data sets, to some extent. A two-step presmoothing method, which was developed to improve the QDC technique results, is also described. Each technique is evaluated by its success at identifying the true QDC of synthetic data sets, which were created to mimic the behavior of real VLF data sets. The best technique is then applied to data sets of real amplitude observations of subionospherically propagated VLF and the results shown for example events. The development of a successful QDC finding algorithm for long-term data sets will allow for the detection of and statistical analysis of the ionospheric response to space weather events, observed through subionospheric VLF.

2. Data Sets

2.1. AARDDVARK VLF Observations

The Antarctic-Arctic Radiation-belt (Dynamic) Deposition-VLF Atmospheric Research Consortium (AARDDVARK)

[Clilverd *et al.*, 2009] is a global network of VLF receivers located primarily in the polar regions (http://www.physics.otago.ac.nz/space/AARDDVARK_homepage.htm). The AARDDVARK receivers providing data for this study are located near Scott Base (SB: 77°50'S, 16°39'E), Antarctica, and Edmonton (EDM: 53°21'N, 112°58'W), Canada. The Scott Base receiver provides examples of long-distance transmitter-receiver great circle paths, while the Edmonton receiver provides examples of relatively short distance paths. Observed signals from short and long paths are expected to behave differently due to the attenuation of higher-level modes with distance [Wait, 1996b]. The data sets chosen from each receiver have the strongest signals of those monitored by those receivers and so provide relatively clear examples. Both receiver installations use UltraMSK software [Clilverd *et al.*, 2009] to process the VLF observations into amplitude and phase data. Figure 1a shows the locations of the two receivers, the monitored VLF communications transmitters that we utilize in our study, and the great circle paths between them.

In the current study we use only the VLF amplitude data, because there are extra difficulties in making long-term phase data sets consistent. An example is unpredictable phase jumps across periods when the transmitters are turned off, as discussed in Rodger *et al.* [2012]. We reduce the 0.2 s resolution raw data to 1 min resolution by averaging; i.e., each data point in the reduced data set is obtained by calculating the median of the 300 data points in that minute of raw data. This averaging removes much of the noisy variation inherent in the received signal from modulation of the transmission. By the law of large numbers, over the course of 1 min this variation has a Gaussian distribution.

For the techniques described in this paper the data was arranged into a matrix such that each row contained 1 day of data (1440 data points), with the rows ordered sequentially in time. Periods of abnormal transmission or interference from the receiver surroundings were removed from the data set. Abnormal transmission periods were identified heuristically as when the signal dropped suddenly to the noise floor between 1 min and the next and then later returned to normal signal levels. Periods of noise interference from the receiver surroundings were identified by comparison to observations on the 23.0 kHz frequency, which is rarely transmitted on.

Consistent temporal spacing of data points is essential for our QDC finding techniques, so periods when data were missing or removed were included as approximating values. The data approximations were done by combining the “linear” method of the “interp1q” MATLAB function for data gaps of <2 h duration, the “TriScatteredInterp” two-dimensional surface interpolation function (using the dimensions of the data matrix) for data gaps 2 h to 2 days, and a median of the same numbered days of data from surrounding years for longer data gaps to maintain the overall coherence of the diurnal pattern within the approximated values.

Figure 1b shows 32 months of amplitude observations of the NDK (25.2 kHz) transmission received at the AARDDVARK antenna located near Edmonton, Canada. The great circle path between transmitter and receiver is completely dark in the primarily red region between 02 and 12 UT and fully sunlit in the green to orange region between 14 and 24 UT. The border between the night and day regions is defined by the twilight modal minima, which vary their time of occurrence regularly through each year according to when the day-night terminator crosses the path.

2.2. Synthetic Data Set Creation

We created synthetic amplitude data sets for the purpose of evaluating the success of our QDC finding techniques at identifying the true underlying QDC of a data set. These synthetic data sets were designed to be representative in their general response to light levels along a propagation path and to space weather events, rather than be a true model of the VLF data set for the equivalent path. The synthetic data set matrices contain 4 years of data at 1 min resolution. Like the AARDDVARK data matrices, each row is 1 day of data arranged in UT time. Background patterns in the synthetic amplitude data simulate the general patterns seen in VLF amplitude data, with periods designated daytime (path fully sunlit), nighttime (path fully dark), and twilight time (day-night terminator along the path).

We present one of our synthetic data sets here, shown in Figure 1c, to illustrate the approach. Figure 1b shows AARDDVARK observations for the equivalent path: NDK-EDM (shown in Figure 1a). The daytime, nighttime, and twilight time of the synthetic data set are defined by the solar zenith angles (SZA) at NDK and EDM. The diurnal variation in the synthetic data set consists of four sections: a constant-valued section representing the VLF response to nighttime conditions (approximately 02–12 UT), a curved section representing the VLF response to daytime conditions (approximately 14–24 UT), and two twilight sections separating the

nighttime and daytime, each with a single sinusoidal minima representing the twilight modal minima as seen in VLF data. The daytime curve (Data_{day}) is calculated as

$$\text{Data}_{\text{day}} = -(\text{SZA}_{\text{NDK}} + \text{SZA}_{\text{EDM}})/2 + 90$$

where SZA_{NDK} and SZA_{EDM} are the SZAs at NDK and EDM, respectively. A long-term trend of a single sinusoidal cycle is imposed on each column in the matrix. The diurnal variation is added to the long-term trend to form the background of the synthetic data set, which is the true QDC that our techniques are aiming to identify. This background forms the dominant variation seen in Figure 1c. Perturbations are imposed, by addition, on the synthetic background to represent the VLF response to solar flares, EEP, and multiday disturbances to the D region. A fourth component imposed on the background represents the effect of random noise on the VLF signal. Figure 1d shows the background and combined data for a representative day from the synthetic data set.

Across the 4 years of our synthetic data set we impose 5000 “EEP events,” which we represent by downward pointing triangles, and 1000 “solar flare” events. The equation used to represent a solar flare event (Flare) is

$$\text{Flare} = 2x \exp(-x/\text{size})$$

where x is minutes from the start of the event and size is a random scale factor from 1 to 20 but biased to the lower end of the range. The imposed EEP events are placed only in the nighttime region of the data set, while the solar flare events are placed only in the daytime region of the data set. The timing of both the EEP and the solar flare events is otherwise random but biased toward periods of geomagnetic disturbance. While these events may not be strictly linked to geomagnetic disturbance, this bias gives a good representation of the clustering of space weather events which occurs in the “real world.” Kp index values for the 4 years spanning January 2009 to December 2012 (sourced from <http://wdc.kugi.kyoto-u.ac.jp>) provide a simple proxy for both solar and auroral activities and are used to supply the bias, where a higher Kp will lead to more imposed synthetic EEP and solar flare perturbations. The magnitude of the imposed EEP and solar flare events is randomly generated within the range 0.6–15 dB, which is representative of the range of responses caused by solar flares and EEP seen in real VLF data sets.

Multiday perturbations are included to simulate the effect of longer space weather events, such as SPEs, or longer-term geomagnetic disturbances. In our synthetic data sets the timing and strengths of these perturbations in the data set are determined by the Dst index values (sourced from <http://wdc.kugi.kyoto-u.ac.jp>) for the same 4 year period as used for the Kp -based perturbations. The range of Dst values in the period was divided into disturbance levels, which were used to assign perturbation values to each entry in the synthetic data matrix. These added values were smoothed to remove sharp steps from the perturbations. The magnitude range of the added values is 0–5 dB, negative during nighttime and positive during daytime. We placed no restrictions on the length of the multiday perturbations, beyond those inherent in the Dst data set disturbance levels.

The added noise component consists of random values selected from a zero-centered Gaussian distribution in the range $\pm x$ that are added to each data point in the daytime and nighttime sections of the data set. We define x from the uncertainties reported by *Rodger et al.* [2007]. The distribution standard deviation during daytime is 0.02 to give an x of 0.1 dB and during nighttime is 0.1 to give an x of 0.5 dB.

3. Technique Descriptions

Below we give descriptions of the QDC finding techniques developed in this study. We also describe the pre-processing addition that we developed to improve the results of the techniques. In the development of these algorithms, we have aimed to make them generic and not specific to one known data set. As such, these approaches should be valid for any subionospheric VLF amplitude data set of sufficient duration.

3.1. Combined Daily Curve

This technique generalizes the method used by *Simon Wedlund et al.* [2014]. They calculated their QDC from the combined curve of several identified quiet days of VLF amplitude observations that occurred shortly before a period of geomagnetic disturbance. In the current study this method is generalized by applying

the technique with no regard for the level of disturbance in the previous days' data; i.e., there is no attempt to determine if the previous days are indeed quiet. This is done so that our technique does not rely on the time-consuming manual identification of quiet days within a data set. We therefore note that the calculated QDC will be of lower quality than if we knew that the utilized observations came from a truly quiet period. Thus, this technique may best suit periods of lower solar activity. We refer to this method as a combined daily curve (CDC).

The CDC is created by averaging data from the 3 days prior to the day of interest. The CDC technique assumes that the diurnal pattern in VLF data changes very little from day to day, except in response to ionospheric perturbations, which the averaging is expected to remove. This assumption is based on examination of diurnal patterns in VLF data sets (e.g., the relatively regular variations seen in Figure 1b).

The CDC is calculated at 10 min resolution, with each value in the CDC being averaged from the same respective 10 data points in each of the previous 3 days. Thus, each average value is calculated from thirty 1 min data points to match the 30 data points, from a single day, that were used by *Simon Wedlund et al.* [2014] for their QDC value calculation. The CDC is then interpolated back to 1 min resolution, using the MATLAB function "interp1" with the "linear" method, for direct comparison with the data. We change the resolution in this manner to reduce the influence of any one data point on the result. The first 3 days of data in the matrix do not have corresponding CDCs as they do not have at least 3 days prior to them.

3.2. Principal Component Analysis

Principal component analysis (PCA) is a tool used in multivariate analysis to expand a data set along its directions of maximal variance. For analysis of data variation, this expansion is sufficient. However, it is possible to summarize the patterns in a data set by selecting expansions along a limited number of directions of highest variance and recombine them [Collier, 2009]. For the purpose of this QDC finding technique, we assume that the majority of the variance in the data set comes from the regular diurnal patterns of the data and is thus concentrated in the lower ordered PCA directions.

The steps of the PCA QDC finding technique for an $m \times n$ data matrix \mathbf{X} are as follows:

1. Create the recentering matrix $\bar{\mathbf{x}}$, which has the entries of each column as the mean of the corresponding column of \mathbf{X} .
2. Calculate the covariance matrix \mathbf{S} of the recentered data matrix.

$$\mathbf{S} = \frac{1}{m-1}(\mathbf{X} - \bar{\mathbf{x}})(\mathbf{X} - \bar{\mathbf{x}})'$$

3. Find the eigenvectors and eigenvalues of \mathbf{S} . These should be sorted in decreasing order by the eigenvalues. The eigenvectors are the directions of maximal variance for the PCA process, and the corresponding eigenvalues give the variance accounted for by each direction.
4. Project the recentered data matrix onto the eigenvectors of \mathbf{S} to find the principal components (PCs). Defining \mathbf{G} as the matrix of eigenvectors, arranged column wise, the matrix of principal components, \mathbf{Y} , is

$$\mathbf{Y} = (\mathbf{X} - \bar{\mathbf{x}})\mathbf{G}.$$

Each column of \mathbf{Y} is a single PC. The PCs are ordered according to the variance accounted for by their corresponding directions, with the first being the projection of the recentered data matrix onto the direction of highest variance.

5. Choose and apply the criteria to be used for limiting the number of PCs. We use the Kaiser criterion [Kaiser, 1960], which retains only those PCs that individually account for more than the mean variance over all the PCs.
6. Invert the projection for all retained PCs, sum them together, and add the recentering matrix. With $\mathbf{y}_{(1,2,\dots,i)}$ and $\mathbf{g}_{(1,2,\dots,i)}$ defined as containing the retained PCs and corresponding eigenvectors, respectively, the resulting QDC matrix \mathbf{Q}_{PCA} is

$$\mathbf{Q}_{\text{PCA}} = \mathbf{y}_{(1,2,\dots,i)}\mathbf{g}'_{(1,2,\dots,i)} + \bar{\mathbf{x}},$$

3.3. Fast Fourier Transform

The fast Fourier transform (FFT) is used to identify the discrete frequency spectrum of a digital data set. In this study the two dimensions of the FFT are the diurnal variation in the rows of the data matrix and the day-to-day variation, which includes the yearly variation, in the columns of the data matrix. Our FFT QDC finding technique uses the two-dimensional transform to calculate the spectrum of a data set, which is then restricted as described below. We calculate the inverse transform of the restricted spectrum to provide our QDC. *Amidor* [2013] gives an overview of the transform in multiple dimensions including details of various issues to be aware of when using the transform.

In this technique we want to remove as much of the perturbation contribution from the spectrum as possible while retaining as much of the background contribution as possible, as this represents the true QDC we are trying to find. The central aspect of this technique is the identification of the spectral components that are dominated by the perturbation spectrum. Once these unwanted components are identified, we remove their contribution to the spectrum by setting them to zero. The QDC is taken as the real component of the resulting matrix from the inverse FFT. Note that providing the spectrum restrictions maintain the symmetry properties of the original spectrum, the result of the inverse FFT will have no imaginary component.

The linear property of the FFT allows for the examination of the features of the synthetic background spectrum independently from the perturbation spectrum. From this examination we are able to identify consistent features of these spectra across multiple synthetic data sets with different backgrounds and thus develop methods to identify perturbation-dominated spectral components for removal from the spectra.

The first spectral restriction is the removal of certain rows of the FFT spectrum to clarify the yearly, including seasonal, variation of the data set. For this clarification to be most effective, the data set is required to be a whole number of years, say p , in length. Cutting the data set prior to application of the FFT may be required to achieve this. The yearly background pattern of a p years length data set repeats p times in the vertical direction of the data matrix. This regular repetition places the background-related spectral components on the p^{th} multiple vertical frequencies, or rows from the center, of the spectrum. Spectral leakage is a frequency-smearing artifact in the FFT that results from the effective discrete truncation of a continuous function [*Amidor*, 2013]. It causes all spectral components in the spectrum to contribute to those surrounding them; in this case the result is that the non- p -multiple rows of the spectrum have some contribution from the background patterns. By limiting the data set to whole numbers of years, we minimize that contribution, allowing us to assume that the non- p -multiple rows are perturbation dominated. Thus, by keeping the data set to p years, we can immediately identify the non- p -multiple rows of the spectrum as being perturbation dominated and set their components to zero for QDC generation.

This first spectral restriction essentially requires data sets to be of longer duration than 2 years to allow for row removal in the spectrum. Due to this requirement, our FFT QDC finding technique is not valid for VLF data sets shorter than 2 years.

The second spectral restriction is the removal of two regions of the spectrum matrix that are consistently perturbation dominated and are located vertically up and down from the center of the matrix and the retention of background-dominated regions. Separate examination of background and perturbation spectra from our synthetic data sets showed us the regions in the combined spectra where each would be expected to be dominant. The strong spectral components of the background layers are located in the center of their spectra, fanning outward horizontally and diagonally with decreasing magnitudes in patterns specific to each background. Figure 2a shows the spectral magnitudes of the central section of the synthetic spectrum. Here the background-related pattern is seen on every fourth row as a higher magnitude than surrounding values. None of the background spectra fan out in the vertical directions. The strongest spectral components of the perturbation layers are located in the central column of their spectra (the vertical green columnar region in Figure 2a), symmetrically reducing in magnitude with horizontal distance. From these observations we find that the two triangular regions located in the vertical directions from the center of the spectrum have little contribution from the background spectra and are thus perturbation dominated. The boundaries of the region of strong background-related spectral components are different for each background and must be identified separately for each data set. Once the boundaries of the region of significant background contribution are identified, the spectral components in the triangular regions outside of the boundaries are easily set to zero using a stencil.

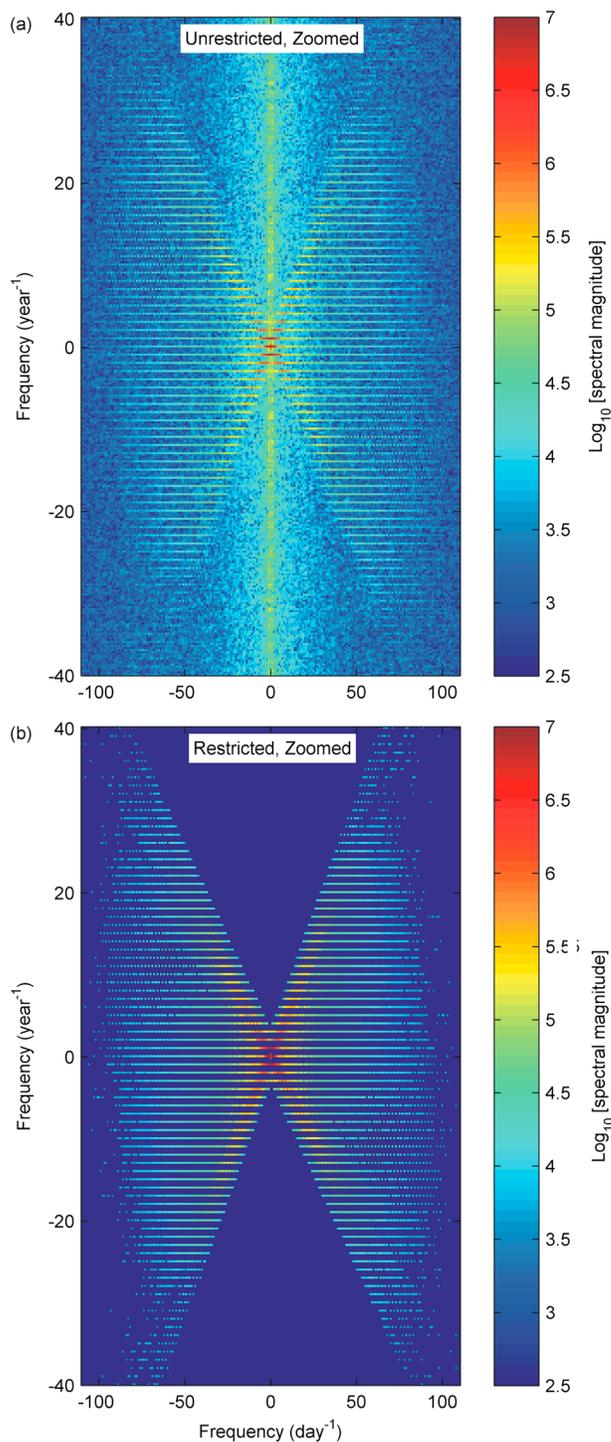


Figure 2. Magnitudes of the two-dimensional FFT spectra for the synthetic data set shown in Figure 1c. (a) Basic spectrum before the restrictions are applied from the FFT QDC finding technique. (b) Fully restricted spectrum. Both plots have been zoomed in to frame the central background-related spectral pattern. The color scale is \log_{10} and shown at the right of each plot.

The third spectral restriction is the removal of low-energy spectral components in the high-frequency regions of the spectrum matrix. At the edge of the spectrum matrix, where the frequencies are highest, the spectral components are perturbation dominated and the spectral magnitudes are relatively low. It is necessary to identify the border of the matrix region within which the background-related spectral components are dominant. This is the point where the distinct pattern of the background-dominated spectral components is subsumed into the general spectrum. A spectral energy limit is employed, with the limit chosen as the lowest energy at which the background pattern is retained and a minimum of spectral components from outside of the pattern are included. This method is less subjective than a determination through visual inspection to find the border of the background-dominated region of the spectrum. The spectral energy limit is different for each spectrum due to the differing background patterns in each corresponding data set. For the chosen energy limit, a plot of the inverse FFT of the discarded spectral components should not include background patterns from the data set or periodic variations of greater than 0.1 dB magnitude.

The first spectral restriction tends to remove contributions from long-term trends to the spectrum of the data set, due to VLF data set long-term trends likely being a response to the solar activity cycle of 11 and 22 years. Unless the data set is itself a multiple of 11 years in length, the main trend-related components are lost at the row removal stage. Thus, for this QDC finding technique to take into account any long-term trends, an extra step is needed to reinclude the strongest of the removed spectral components in the low-frequency region of the matrix to the spectrum prior to the inverse transform.

The final synthetic spectrum, after all the restrictions have been applied, is shown in Figure 2b. As with Figure 2a we show only the spectral magnitudes from the central section of the spectrum. The combination of the row removal and stencil restrictions has removed the visible contribution of the

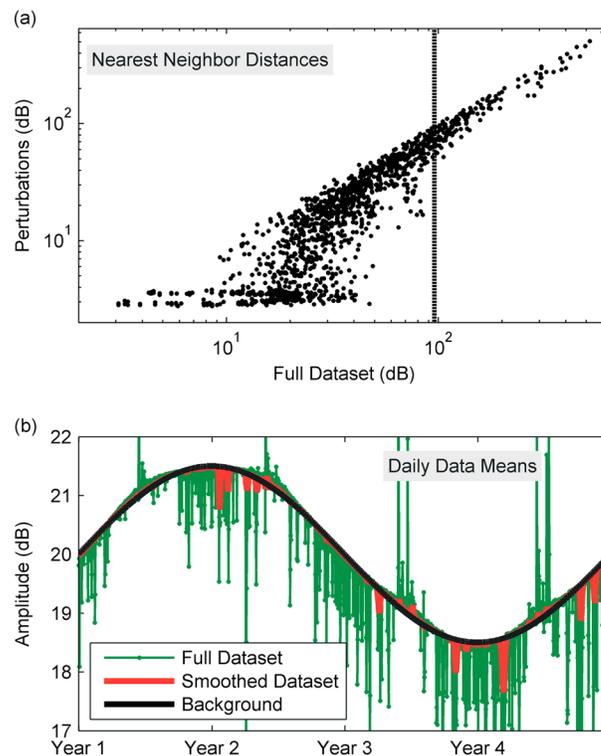


Figure 3. (a) Nearest-neighbor distances between rows of the synthetic perturbation matrix versus the distances between the corresponding rows of the full synthetic data set (perturbations and background combined). Vertical dashed line indicates top 10% of full data distances. (b) Daily means for the full synthetic data set (green line and markers), smoothed data set (red line), and background of the data set (black line).

The results of all three QDC finding techniques are negatively influenced by periods of significant disturbance in the data sets to some degree. We investigated nearest-neighbor distances [Cover and Hart, 1967] as a method of defining the disturbance level of a row of data. Figure 3a shows the nearest-neighbor distance for each row of the synthetic perturbation matrix plotted against the nearest-neighbor distance for the corresponding rows of the full synthetic data set. Here we see that rows with higher data set distances also have higher perturbation distances. From this relationship, we determine that the data set nearest-neighbor distance of a row is a good indicator for the actual disturbance level of a row. We therefore remove from the data matrix those rows with the highest 10% nearest-neighbor distances, as the most disturbed. In Figure 3a this limit is marked by a dashed vertical line.

We then smooth the data, which serves two purposes: to replace the removed data from disturbed days and reduce the influence of short-term perturbations, i.e., solar flares, on the QDC. We use the “rloess” method of the “smooth” function from the MATLAB® software package’s Curve Fitting Toolbox. This method is a “local regression using weighted linear least squares and a second degree polynomial model” that “assigns lower weight to outliers in the regression” (www.mathworks.com/help/curvefit/smooth.html). The rloess method was preferred for the smoothing over a moving average, because of the lower influence of outlying values on the result under this method. This smoothing method fills gaps in the input data as part of the algorithm. We found that smoothing over the gaps from the removed disturbed days in the data matrix improves our results even more than filling them with representative values. The long-term trend in the data is not significantly affected by this method of smoothing as shown by the daily data means presented in Figure 3b. The smoothed data set daily means show significantly less variation than those of the unsmoothed, full, data set while also remaining close to the background daily means.

The smoothing is done both column wise and row wise in the data matrix. The column-wise smoothing is intended to remove single-day perturbations, which can be considered outliers within the general shape of the data from day to day, and mitigate the effect of multiday perturbations, such as SPEs. The row-wise smoothing is intended to further reduce the effect of noise around the signal.

perturbation-dominated components in the central region of the spectrum, while the removal of lower energy components shows the border of the background-dominated matrix region.

3.4. Additional Smoothing

As will be reported in section 4.2, the three basic QDC finding techniques, described above, produce promising results when applied to our synthetic data sets. We also investigated methods to presmooth the data sets with the aim of improving the results from the basic techniques. We found that a two-step preprocessing approach, which involves the removal of the most disturbed days of data and then a smoothing of the resulting matrix, applied to the data set prior to application of the QDC finding technique provided an improvement in the results for the daytime and nighttime regions of the matrix. These preprocessing methods are described below.

The results of all three QDC finding techniques are negatively influenced by periods of significant disturbance in the data sets to some degree. We investigated nearest-neighbor distances [Cover and Hart, 1967] as a method of defining the disturbance level of a row of data.

Figure 3a shows the nearest-neighbor distance for each row of the synthetic perturbation matrix plotted against the nearest-neighbor distance for the corresponding rows of the full synthetic data set.

Care must be taken in choosing the span for the smoothing. Too high a span and the desired background patterns in the data are lost, too low and the smoothing is practically pointless. We tested a range of spans on various of our synthetic data sets to determine the level required under these constraints. For the twilight times, we found that a span of 7 data points provides adequate smoothing of perturbations without significantly altering the shape of the minima. A higher span is possible for the daytime and nighttime regions of the data matrix. We found that a span of 13 data points provided very good smoothing while limiting the addition of negative artifacts to the smoothed data matrix in these regions. We therefore smooth the data matrix twice, once at a span of 7 and once at 13 and combine the twilight time region of the 7-span result with the day and nighttime regions of the 13-span result to give our final smoothed data set for application of a QDC finding technique.

A low-pass filter might be used here as an alternative to the smoothing. However, it is not clear whether this style of filter would provide a significant enough improvement to the results of the method described above to justify the added subjectivity of determining the cutoff frequency for each data set. Our smoothing method is convenient to the MATLAB[®] user and requires little subjectivity in the identification of the required span, which can then be easily translated across different data sets.

4. Testing Techniques on Synthetic Data

4.1. Method to Quantify Technique Success

We evaluate the success of our QDC techniques by calculating a parameter to indicate how close our QDC matrices are to the synthetic background, which is the true QDC of the synthetic data set. This parameter allows us to directly compare the success of our techniques. We calculate this parameter from the difference between the QDC and the background, which we refer to as the Comparison. Clearly, it is only possible to determine this parameter for synthetic data sets due to the true background being unknown for real VLF observations.

Our indicative parameter is based on the L^2 vector norm, and so we will refer to it as the norm for the remainder of this study. The equation used to define the norm is

$$||\mathbf{v}|| = \sqrt{\sum_i v_i^2 / n} \quad (1)$$

where $||\mathbf{v}||$ is the norm, v_i are the entries in the relevant section of the Comparison matrix, and n is the number of entries in the section. The norm parameter is higher than a simple average of absolute values due to the squaring of the entries. It has no direct physical meaning, being used here as an estimation of the outer variability of the Comparison matrix. The norm can be calculated for each section of the Comparison matrix, nighttime, daytime, and twilight time, as well as for the complete matrix. This allows us to compare technique success between Comparison sections.

For our technique evaluation we use 10 different synthetic data sets, with identical backgrounds, that differ only in the random timing and magnitude of the imposed perturbations. The final reported norms, in Table 1, for each technique are the mean of the 10 norms found for the application of the specific technique to each of the 10 data sets. The uncertainty is taken as the range of the norms over the 10 data sets and is also reported in Table 1.

Table 1 has two sections, with the norms of the upper section for application of the QDC finding techniques (outlined in sections 3.1–3.3) to the synthetic data, and the norms of the lower section for the inclusion of the two-step presmoothing method (outlined in section 3.4) prior to application of the techniques. The norms in each section of the table are arranged by technique and region of the Comparison matrix: “All” for the entire synthetic data set, “Day” for periods when the path is fully sunlit, “Night” for the periods when the path is fully dark, and “Twilight” for the periods when the day-night terminator intersects the path.

Lower norms result from technique-calculated QDCs that are closer to the synthetic background, on average. Thus, the best technique is the one resulting in the lowest norms. The norms in the top row of Table 1 compare the complete synthetic data set, including all the imposed perturbations, to its background. These norms are the absolute upper boundary of what we would accept for the results from a QDC technique as a higher norm would imply that extra perturbations have been added by a technique.

Table 1. Norms (Equation (1)) for the Comparison of Our Calculated QDCs to the True QDCs of Our Synthetic Data Sets^a

	All	Day	Night	Twilight
Data	2.65 ± 0.08	1.71 ± 0.04	3.45 ± 0.14	1.74 ± 0.15
CDC	1.48 ± 0.03	1.24 ± 0.01	1.33 ± 0.08	2.86 ± 0.01
PCA	2.53 ± 0.08	1.59 ± 0.04	3.24 ± 0.14	2.36 ± 0.10
FFT	0.85 ± 0.01	0.82 ± 0.01	0.9 ± 0.02	0.68 ± 0.01
Smoothed	0.57 ± 0.03	0.44 ± 0.05	0.24 ± 0.02	1.61 ± 0.14
CDC	0.96 ± 0.02	0.64 ± 0.02	0.29 ± 0.01	2.95 ± 0.05
PCA	0.79 ± 0.04	0.54 ± 0.03	0.47 ± 0.02	2.17 ± 0.12
FFT	0.28 ± 0.01	0.23 ± 0.02	0.14 ± 0.01	0.77 ± 0.05

^a All values are rounded to two decimal points. Units are dB.

The following subsections give the quantitative evaluation of the “Basic” and “Presmoothing” techniques by their norms, as displayed in Table 1. Qualitative evaluation is provided for each technique by Comparison plots, i.e., the difference between the calculated and true QDCs. These plots are given in Figures 4 and 5 and are each processed from the same representative data set of the 10 used in the testing.

4.2. Evaluation of Basic Techniques

Comparing the norms within the upper section of the table, we see that in the twilight sectors the CDC and PCA QDCs result in higher norms than those for the synthetic data set itself. In the day and night sectors, all three QDC techniques result in lower norms than those of the data set. Across all sectors the FFT QDC finding technique shows the best results, with norms of less than 1 dB, whereas the CDC and PCA techniques both result in norms greater than 1 dB.

Figure 4 shows the Comparison plots for the synthetic data and all three basic QDC techniques. Figure 4a, data, is effectively just showing the synthetic perturbations, as expected. Figures 4b and 4c, for the basic CDC and basic PCA techniques, respectively, show significant remaining influence of the imposed multiday perturbations. Figure 4d, basic FFT, shows less localized influence of the perturbations than is seen in the plots for the other techniques. However, the overall effect of the imposed perturbations for this technique is to bias the calculated QDC in the dominant direction of the data disturbance, either positive (daytime for the synthetic data sets) or negative (nighttime for the synthetic data sets).

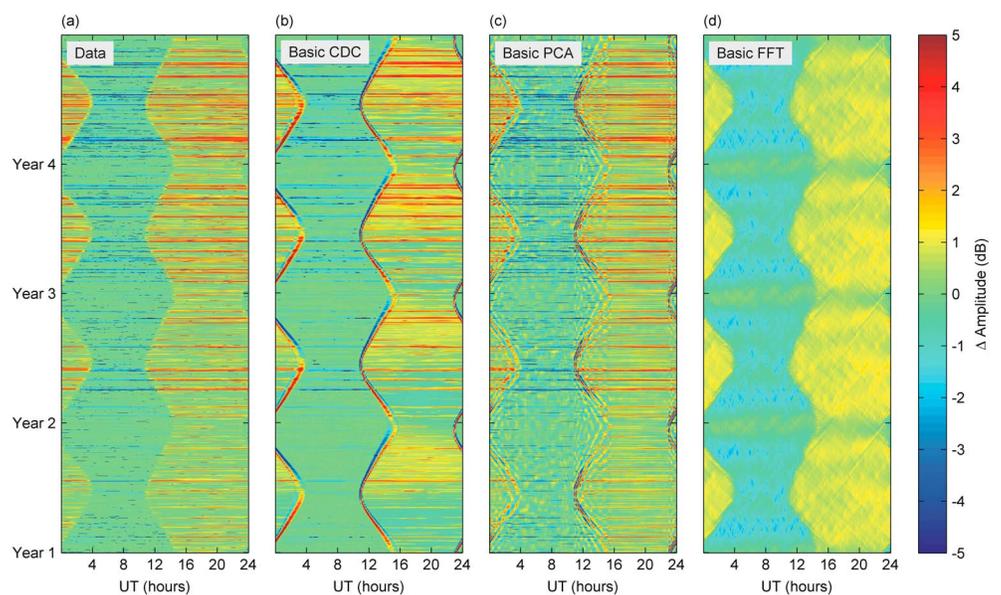


Figure 4. Comparison matrices, i.e., the difference between the calculated and true QDCs, for the full synthetic data set and three QDC finding techniques. The technique used to calculate the corresponding QDC is given in the top left of each plot. All plots are on the same color scale, which is shown at the right of the plots.

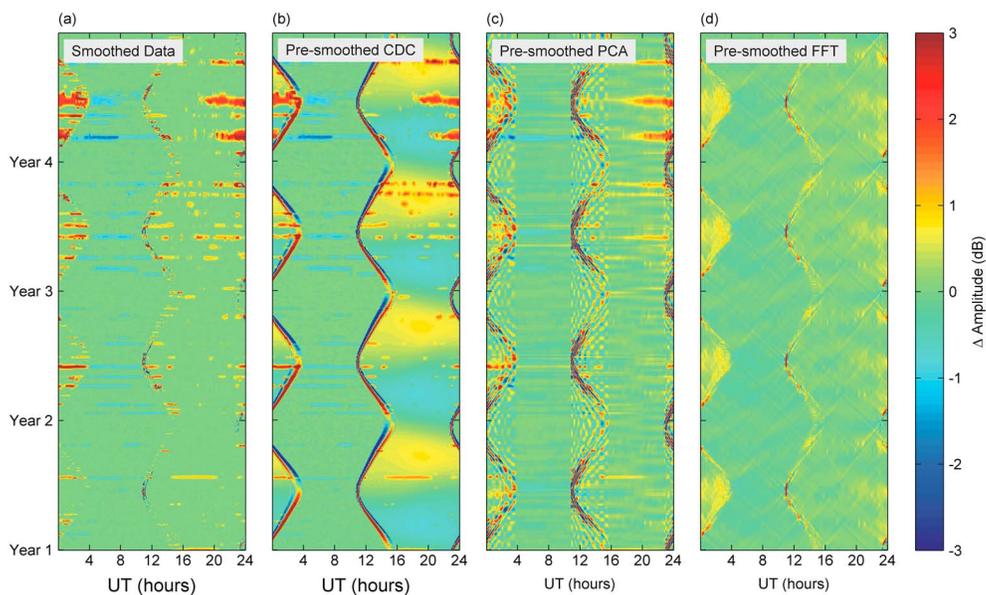


Figure 5. Comparison matrices for the smoothed synthetic data set and subsequent application of the three QDC finding techniques. The technique used to calculate the corresponding QDC is given in the top left of each plot. All plots are on the same color scale, which is shown at the right of the plots and is smaller than that of Figure 4.

The norms and Comparison plots for the three basic QDC finding techniques indicate that the basic FFT technique is promising but has the significant issue of bias, which will be important in practical application. However, further investigations found that these results can be significantly improved upon, and the next section gives the analysis for the addition of the developed presmoothing method to the techniques.

4.3. Evaluation of Presmoothing Techniques

The complete algorithm for each technique evaluated in this subsection involves applying the two-step presmoothing method, described in section 3.4, to the full synthetic data set and then applying the chosen QDC finding technique to the resulting data matrix.

The top row of the lower section of Table 1 gives the norms for the Comparison of the smoothed synthetic data set to the background. Here we see an immediate improvement over all of the basic norms in the upper section of the table, excepting only the twilight norm for the FFT technique.

Applying either of the CDC and PCA techniques to the smoothed synthetic data gives no improvement to the norms over the smoothing alone. Applying the FFT technique to the smoothed data improves the results in all sectors, almost halving the norms from the smoothing alone. The daytime norm for the presmoothed FFT technique is 0.23 dB, which is around twice the maximum level of the daytime imposed noise (0.1 dB). The nighttime norm is 0.14 dB, which is less than half of the maximum level of the imposed nighttime noise (0.5 dB). In contrast, the norm for the twilight time section has increased compared with that of the basic FFT technique.

Figure 5 shows the Comparison plots for the smoothing method and presmoothed QDC finding techniques. Note that the color scale range of this figure has been decreased from Figure 4. Figure 5a, smoothed data, shows significant removal of perturbations from the calculated QDC, with only localized influence of highly perturbed periods in the synthetic data set. Figure 5b, presmoothed CDC, shows no improvement over the smoothed plot during the times of highly perturbed periods. The yellow and blue regions between 12 and 24 UT in the CDC plot show that a simple average of previous days as a QDC is prone to influence from any day-to-day slope present in the data, i.e., during sunlit periods in the synthetic data set (14–24 UT, Figure 1c). Figure 5c, presmoothed PCA, shows the difficulty of separating background-related variance from perturbation-related variance in the PCA process. In this plot, the vertical sections encompassing the periods of twilight modal minima (23–04 UT and 11–16 UT) show a distinct lack of definition for the minima while other sections are clearly influenced by the perturbations remaining in the smoothed synthetic data, such that they appear in our calculated QDC. The PCA QDC finding technique may have more success at identifying

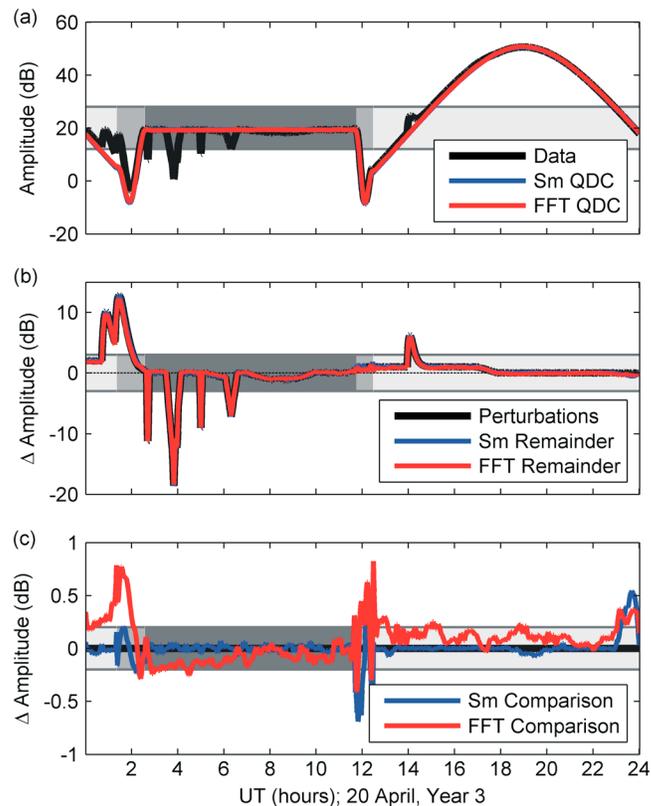


Figure 6. (a) Synthetic data for 1 day (black line) and the calculated QDCs found by the smoothing process (blue line) and FFT technique (red line). (b) Perturbations in the data set and the Remainders from the techniques. (c) Comparisons between the calculated and true QDCs. All three plots have a guide bar as to the level of light on the path, either fully sunlit (light grey), fully dark (dark grey), or mixed with the terminator located across the path (mid grey). The date of the day is given in the x axis label to allow cross checking with Figure 1c.

the true QDC for a shorter-period data set, of maybe month duration; however, investigation of this possibility is beyond the scope of this study. Figure 5d, presmoothed FFT, still shows some bias in the calculated QDC to the dominant direction of the data disturbance; however, this bias has been significantly reduced from that seen in the basic FFT Comparison plot of Figure 4. While the presmoothed FFT technique does not represent the modal minimum periods well, in general, this technique provides the best calculated QDCs.

We conclude from the norms presented in Table 1, and examination of the plots in Figures 4 and 5, that the best of the methods considered in this study for identifying a QDC of a long-lasting VLF data set is to smooth the data set as described in section 3.4 then apply the FFT technique as described in section 3.3. Unfortunately, the restriction of the FFT technique to data sets of at least 2 years duration, to allow the row removal step to be applied, means that this technique is not appropriate for shorter data sets. Thus, for data sets of less than 2 years duration we recommend the presmoothing process alone as the best method for identifying a QDC.

Figure 6 shows a single representative day of synthetic data and the results for the presmoothing process and the FFT QDC finding techniques. Figure 6a is the synthetic data and the two QDC results, which follow the diurnal pattern in the data visually successfully. Figure 6b shows the imposed perturbations for the day and the difference between the data and each QDC, which we call the Remainder. At the visual level, the Remainders contain the imposed perturbations. Figure 6c shows the Comparison, which is the difference between the true and calculated QDCs or equivalently between the perturbations and the Remainder, for the two QDC results. For this day, the FFT has larger magnitude Comparison values than the presmoothing process does, in general; however both lines on the plot remain within 0.25 dB of zero for most of the day.

We note that while the described methods give good results for identifying the QDC from perturbations occurring during relatively slowly changing sections of data, such as is usually seen when the VLF path is either fully sunlit or fully dark, the sharp amplitude changes seen around the twilight modal minima times are not

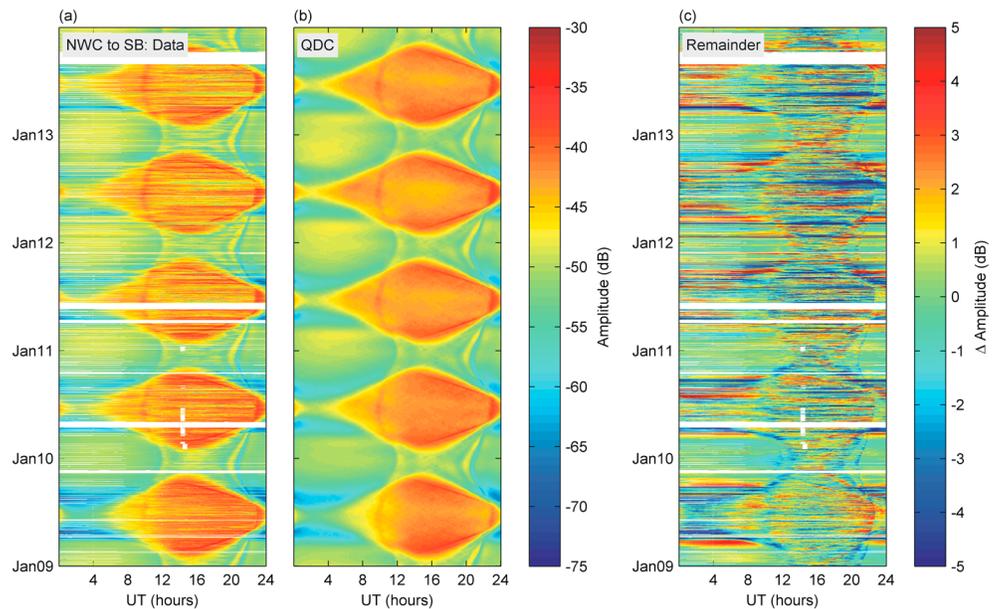


Figure 7. (a) Observations of the NWC transmission received at the SB antenna from January 2009 to December 2013. White areas of the plot show the placeholder values replacing unusable data. (b) The QDC calculated using the presmoothed FFT technique. The color scale for the data set and QDC plots is given at the right of the QDC plot. (c) The Remainder or difference between the data set and the calculated QDC, with color scale at the right of the plot.

so well dealt with. At this point we struggle to produce an accurate QDC representing the intensity of twilight time amplitude variations. Therefore, caution is advised in the interpretation of QDC finding technique results around the times of twilight modal minima.

5. Application to Actual AARDDVARK Data Sets

We now provide example results of the application of this overall technique to our AARDDVARK VLF data sets. We take the smoothing spans that were used for the synthetic data sets and use these spans for the smoothing of the AARDDVARK VLF data sets.

5.1. Clarifying the FFT Spectrum

When we began applying our FFT QDC finding technique to real VLF observations, we found that the background-dominated central pattern of the FFT spectrum was less distinct for some data sets than for the synthetic data set. This lack of clarity of the central pattern was identified as being caused by two sources. First, the dynamic range of amplitudes for a VLF data set is usually much less, varying from 42 to 55 dB for the data sets used in this study, than the approximately 100 dB used for the synthetic data set. That value was set to ensure clear diurnal variations rather than as an actual model of real VLF data. Second, the twilight time modal minima patterns in the synthetic data set were based on a relatively short transmitter-receiver path (NDK to EDM in Figure 1a at 1.304 Mm) and so had a very simple structure, which made the background-related spectral patterns clear in the overall spectrum. Longer paths demonstrate more complex twilight modal interference patterns due to there being more distance along the path for interference fringes to occur [Clilverd *et al.*, 1999]. The background-related spectral patterns in the spectrum are less clear as the path lengthens, such as for the three Scott Base recorded transmitters in this study.

In order for the parameters of the restriction stencil to be correctly identified when the amplitude dynamic range is small and the modal interference patterns in the data set complex, the central pattern of the real VLF spectrum needs to be clarified. We do this by subtracting an average magnitude row (found from the perturbation-dominated higher-frequency region of the spectrum) from the magnitudes of each row of the overall spectrum, which leaves an approximate indication of the background-related pattern in the spectrum for identification of the stencil boundaries. The stencil is then applied to the unclarified spectrum as

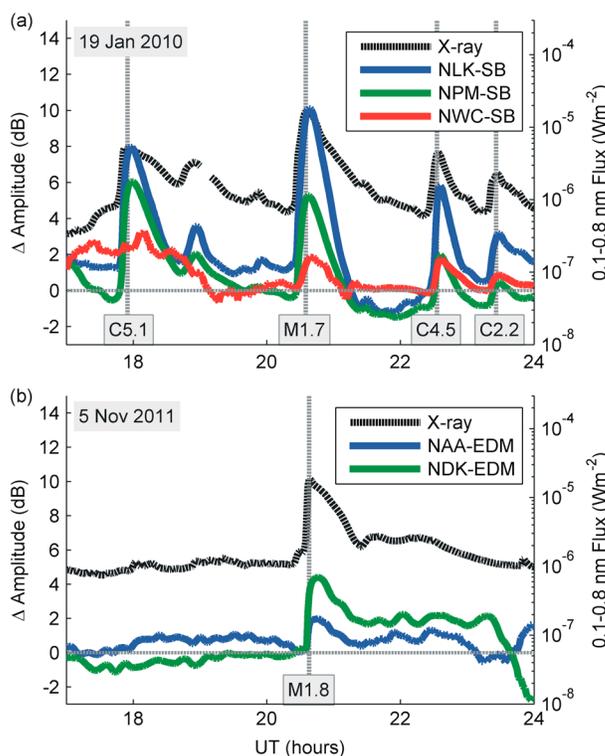


Figure 8. (a) Remainders (observed amplitudes-calculated QDC) for three transmitter signals observed by the Scott Base receiver (solid colored lines, left y axis) for 17–24 UT on 19 January 2010. (b) Remainders for two transmitter signals observed by the Edmonton receiver for 17–24 UT on 5 November 2011. Included on the plots are solar X-ray observations (thick dashed black line, right y axis) from the GOES 14 satellite and the GOES 15 satellite, respectively. Grey dashed horizontal line indicates 0 dB Remainder, i.e., where the calculated QDC equals the data. Grey dashed vertical lines indicate the peak flux times for NOAA-identified solar flares, with the magnitude of each flare given at the base of each line.

visible response to the M1.7 flare, which occurs during the period of partial illumination. The other four paths were fully sunlit during the times of the shown solar flares. Variations in the solar X-ray observations outside of the flares are also seen as variations in the NLK-SB and NPM-SB observations. These examples demonstrate our QDC finding technique’s success at identifying the underlying variation for relatively short duration space weather events.

Figure 9 shows an example of a VLF response to a SPE for the NLK transmission observed by the Scott Base receiver. An SPE is defined for space weather purposes by the proton flux at energies >10 MeV exceeding a threshold of $10 \text{ (cm}^2 \text{ s sr)}^{-1}$ at geosynchronous orbit. The QDC in Figure 9a shows a consistent diurnal variation, which the amplitude data largely follow before the SPE begins and after the SPE flux has returned to relatively quiet levels, i.e., approximately 80–96 h in the plot. Figure 9b shows only the daytime and nighttime Remainders. We do not show the Remainder for the twilight modal minima periods in accordance with the caution advised for the interpretation of the QDC during these periods. The daytime Remainder shows a clear offset from zero for the first two periods when the VLF path is sunlit after the SPE begins. The nighttime Remainder shows a general offset from zero for the first three periods after the SPE begins, although with more variability than the daytime periods show. Note that the SPE is clearly still affecting the data in the third nighttime period even though the SPE flux is below the SPE threshold for this period. Figure 9c shows the corrected >10 MeV proton flux observations from GOES 13 for context. The VLF amplitude response to changes in waveguide parameters varies depending on the result of the superposition of multiple propagating modes. This will not generally lead to a linear relationship between the perturbing SPE flux and the observed Remainder, as this figure shows. The Remainder here demonstrates our QDC finding technique’s success at identifying

normal. With this addition to the FFT QDC finding technique, the response of the real VLF data sets to the technique improves.

5.2. Application Results

Figure 7 shows the data set, calculated QDC, and the difference between the two (Remainder) for 5 years of amplitude observations for the NWC (19.8 kHz) transmission received by the AARDDVARK antenna near Scott Base, Antarctica. The overall background patterns of the data set appear well reproduced in the QDC. However, as the true QDC for real VLF amplitude observations is unknown, this is impossible to quantify. Some of the modal minima regions of the Remainder plot still show consistent amplitude differences, in contrast to the daytime and nighttime regions, where the differences appear dominated by true perturbations.

Details from the Remainders in Figures 7c and 6b suggest that our FFT QDC finding technique is successful at identifying VLF responses to solar flares. This is confirmed for real VLF observations by examples of the VLF Remainder response to solar flares shown in Figure 8. These plots also show the, flare-defining, GOES satellite-observed solar X-ray (0.1–0.8 nm) flux for the same period. The NWC-SB path was partly lit until approximately 21:30 UT when it became fully sunlit but still shows a visi-

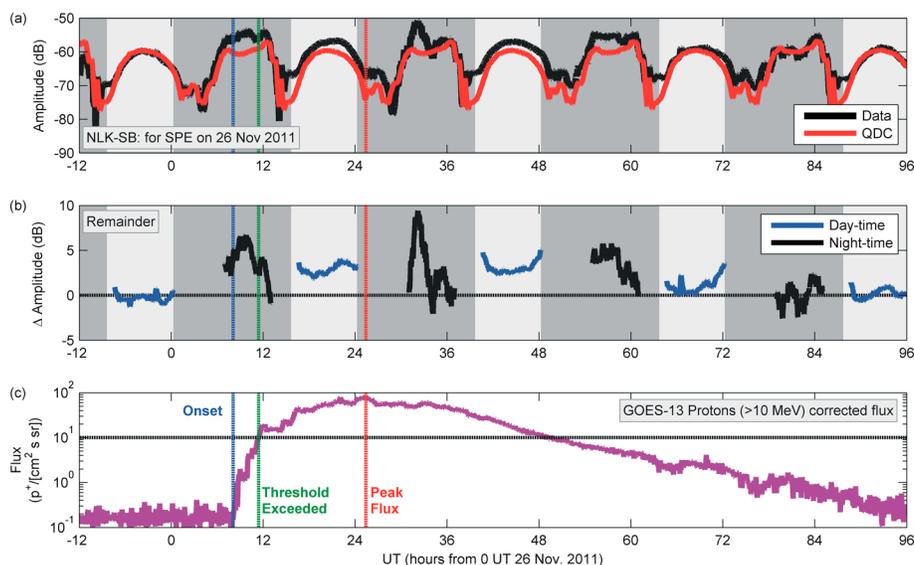


Figure 9. (a) Data from the NLK transmission observed by the Scott Base receiver (black line) and the calculated QDC (red line) for the period of a SPE starting 26 November 2011. The background color indicates the level of light on the path, either fully sunlit (light grey) or with the terminator located across the path (mid grey). (b) Remainder during periods when the path is fully sunlit or mostly dark, with the background color indicating the light level. (c) Corrected >10 MeV proton flux observations from GOES 13. The y axis of this plot is a log₁₀ scale. The threshold for SPE recognition is marked by a horizontal dashed black line. In all plots the dashed vertical blue line indicates the time of onset of the initial flux increase, the green line the time when the SPE threshold was exceeded, and the red line the time of peak proton flux.

the underlying variation even during space weather events lasting multiple days. This figure also shows that the *D* region exhibits sensitivity to solar protons for fluxes below the SPE threshold.

6. Summary and Conclusions

In this paper we described three algorithmic techniques for the calculation of quiet day curves for observations of VLF transmissions propagated subionospherically.

1. The combined daily curve technique calculated an average of the previous 3 days data for its QDC.
2. The principal component analysis technique transformed the data matrix to the directions of maximal variance, selected those directions accounting for more than the mean variance, and transformed them back to data space for its QDC.
3. The fast Fourier transform technique transformed the data matrix to its discrete spectrum, restricted those spectral components likely to be perturbation dominated, and transformed the restricted spectrum back to data space for its QDC.

In addition, a smoothing process was described for application to the data prior to a QDC finding technique.

We evaluated the success of these techniques at identifying the true QDCs of perturbed synthetic data sets and identified the algorithm combining the presmoothing process (described in section 3.4) and the fast Fourier transform-based QDC finding technique (section 3.3) as the most successful technique on average over an entire data set. This combined technique was found to identify the true QDC of our synthetic data sets to within 0.23 ± 0.02 dB during day-defined periods and within $0.14 \pm .01$ dB during night-defined periods of the data sets. The fast modal variations during the twilight-defined periods were identified to within $0.77 \pm .05$ dB. The FFT-based technique is only valid for data sets of at least 2 years; for shorter data sets the presmoothing process alone, which was found to give the second best results in the evaluation, is recommended as a QDC finding technique.

The combined presmoothing and FFT-based QDC finding technique was then applied to real data sets of observed VLF transmissions, from the AARDDVARK receivers located near Scott Base and Edmonton. Example results for five transmitter-receiver paths were provided to demonstrate the technique's ability to identify

responses to perturbations across the entire data set (Figure 7), to solar flares (Figure 8), and to a multiple day SPE in real-world VLF data (Figure 9). From these examples we deduce that this FFT-based QDC finding technique will allow for statistical analysis of VLF responses to space weather events occurring in data sets of longer duration than 2 years.

Acknowledgments

The synthetic data set used throughout this paper and the file used to create it are available from the corresponding author in MATLAB .mat and .m formats, respectively. AARRDVARK VLF data availability is described at its website http://www.physics.otago.ac.nz/space/AARRDVARK_homepage.htm. The GOES 13 proton (>10 MeV) corrected flux data used in Figure 9c was downloaded (2 December 2014) from online file http://satdat.ngdc.noaa.gov/sem/goes/data/new_avg/2011/11/goes13/csv/g13_epead_cpflux_5m_20111101_20111130.csv. The GOES 14 X-ray data used in Figure 8a were downloaded (28 June 2014) from online file http://satdat.ngdc.noaa.gov/sem/goes/data/new_avg/2010/01/goes14/csv/g14_xrs_1m_20100101_20100131.csv. The GOES 15 X-ray data used in Figure 8b were downloaded (12 November 2014) from online file http://satdat.ngdc.noaa.gov/sem/goes/data/new_avg/2010/01/goes14/csv/g15_xrs_1m_20111101_20111130.csv. K.C.M. and C.J.R. would like to acknowledge support from the New Zealand Marsden Fund. K.C.M. would like to acknowledge support from a University of Otago Publishing Bursary. M.A.C. would like to acknowledge support from the Natural Environmental Research Council grant NE/J008125/1.

References

- Amidor, I. (2013), *Mastering the Discrete Fourier Transform in One Two or Several Dimensions: Pitfalls and Artifacts*, Springer, London.
- Barr, R., D. L. Jones, and C. J. Rodger (2000), ELF and VLF radio waves, *J. Atmos. Sol. Terr. Phys.*, *62*(17-18), 1689–1718, doi:10.1016/S1364-6826(00)00121-8.
- Cliilverd, M. A., N. R. Thomson, and C. J. Rodger (1999), Sunrise effects on VLF signals propagating over a long north-south path, *Radio Sci.*, *34*(4), 939–948, doi:10.1029/1999RS900052.
- Cliilverd, M. A., et al. (2009), Remote sensing space weather events: Antarctic-Arctic Radiation-belt (Dynamic) Deposition-VLF Atmospheric Research Consortium network, *Space Weather*, *7*, S04001, doi:10.1029/2008SW000412.
- Collier, A. B. (2009), Principal component analysis of sub-ionospheric propagation conditions, *paper presented at IET 11th International Conference on Ionospheric Radio Systems and Techniques (IRST 2009)*, pp. 86–90, Edinburgh, U. K., doi:10.1049/cp.2009.0041, 28–30 April.
- Cover, T. M., and P. E. Hart (1967), Nearest neighbor pattern classification, *IEEE Trans. Inf. Theory*, *13*(1), 21–27, doi:10.1109/TIT.1967.1053964.
- Han, F., and S. A. Cummer (2010a), Midlatitude nighttime D region ionosphere variability on hourly to monthly time scales, *J. Geophys. Res.*, *115*, A09323, doi:10.1029/2010JA015437.
- Han, F., and S. A. Cummer (2010b), Midlatitude daytime D region ionosphere variations measured from radio atmospheric, *J. Geophys. Res.*, *115*, A10314, doi:10.1029/2010JA015715.
- Kaiser, H. F. (1960), The application of electronic computers to factor analysis, *Educational Psychological Meas.*, *20*(1), 141–151, doi:10.1177/001316446002000116.
- Lynne, K. J. W. (2010), VLF waveguide propagation: The basics, *AIP Conf. Proc.*, *1286*(1), pp. 3–41, Propagation effects of very low frequency radio waves: Proceedings of the 1st International Conference on Science with Very Low Frequency Radio Waves: Theory and Observations, doi:10.1063/1.3512893.
- Rodger, C. J., M. A. Cliilverd, N. R. Thomson, R. J. Gamble, A. Seppälä, E. Turunen, N. P. Meredith, M. Parrot, J. A. Sauvaud, and J.-J. Berthelier (2007), Radiation belt electron precipitation into the atmosphere: Recovery from a geomagnetic storm, *J. Geophys. Res.*, *112*, A11307, doi:10.1029/2007JA012383.
- Rodger, C. J., M. A. Cliilverd, A. J. Kavanagh, C. E. J. Watt, P. T. Verronen, and T. Raita (2012), Contrasting the responses of three different ground-based instruments to energetic electron precipitation, *Radio Sci.*, *47*, RS2021, doi:10.1029/2011RS004971.
- Shao, X.-M., E. H. Lay, and A. R. Jacobson (2013), Reduction of electron density in the night-time lower ionosphere in response to a thunderstorm, *Nat. Geosci.*, *6*(1), 29–33, doi:10.1038/ngeo1668.
- Simon Wedlund, M., M. A. Cliilverd, C. J. Rodger, K. Cresswell-Moorcock, N. Cobbett, P. Breen, D. Danskin, E. Spanswick, and J. V. Rodriguez (2014), A statistical approach to determining energetic outer radiation belt electron precipitation fluxes, *J. Geophys. Res.*, *119*, A019715, doi:10.1002/2013JA019715.
- Thomson, N. R. (1993), Experimental daytime VLF ionospheric parameters, *J. Atmos. Sol. Terr. Phys.*, *55*(2), 173–184, doi:10.1016/0021-9169(93)90122-F.
- Thomson, N. R., M. A. Cliilverd, and W. M. McRae (2007), Nighttime ionospheric D region parameters from VLF phase and amplitude, *J. Geophys. Res.*, *112*, A07304, doi:10.1029/2007JA012271.
- Wait, J. R. (1996a), *Electromagnetic Waves in Stratified Media*, reissued ed., chap. 9: V.L.F. Propagation—Theory and Experiment, IEEE/Oxford Univ. Press, New Jersey.
- Wait, J. R. (1996b), *Electromagnetic Waves in Stratified Media*, reissued ed., chap. 6: Fundamentals of Mode Theory of Wave Propagation, IEEE/Oxford Univ. Press, N. J.
- Wautelet, G., and R. Warnant (2012), Local climatological modeling of ionospheric irregularities detected by GPS in the mid-latitude region, *J. Atmos. Sol. Terr. Phys.*, *89*, 132–143, doi:10.1016/j.jastp.2012.08.015.