Radio Science

RESEARCH ARTICLE
10.1029/2023RS007834

Processing of VLF Amplitude Measurements: Deduction of a Quiet Time Seasonal Variation

H. Schneider[1], V. Wendt[1], D. Banys[1], M. Clilverd[2], and T. Raita[3]

[1]German Aerospace Centre, Solar Terrestrial Institute, Neustrelitz, Germany, [2]Physical Sciences Division, British Antarctic Survey, Cambridge, UK, [3]Sodankylä Geophysical Observatory, University of Oulu, Oulu, Finland

Key Points:

- VLF signal processing with robust statistical methods, Median of Absolute Deviations and Pruned Exact Linear Time
- Deduction of quiet time amplitude behavior through change point detection and systematic outlier filtering
- VLF signal amplitude seasonal variation identified for five different transmitter-receiver paths during undisturbed conditions

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

H. Schneider,
helen.schneider@dlr.de

Citation:

Schneider, H., Wendt, V., Banys, D., Clilverd, M., & Raita, T. (2024). Processing of VLF amplitude measurements: Deduction of a quiet time seasonal variation. *Radio Science*, *59*, e2023RS007834. https://doi.org/10.1029/2023RS007834

Received 2 AUG 2023
Accepted 28 JAN 2024

Author Contributions:

Data curation: V. Wendt, M. Clilverd, T. Raita
Supervision: V. Wendt
Writing – review & editing: V. Wendt, D. Banys, M. Clilverd, T. Raita

Abstract The amplitude of Very Low Frequency (VLF) transmissions propagating from transmitter to receiver between the Earth's surface and the ionospheric D-region is a useful measurement to detect changes in the ionization within the D-region ranging from 60 to 90 km. The VLF signal amplitude is disturbed by geomagnetic, solar, and atmospheric phenomena. To be able to identify perturbations in the VLF signal amplitude, we determine its averaged seasonal variation under quiet solar and geomagnetic conditions. Here it is challenging, that long time series of the VLF signal amplitude show significant jumps and outliers, which are caused artificially by technical adjustments/maintenance work. This paper presents a new approach for processing long VLF data time series over multiple years resulting in level 2 data. The new level 2 data enables the consideration of time series with artificial jumps since the jumps are leveled. Moreover, the outliers are removed by a robust and systematic 2-step outlier filtering. The average seasonal and diurnal variation for different transmitter-receiver combinations can be computed with the new level 2 data by applying a composite analysis. A subsequently applied polynomial fit obtains the quiet time lines for daytime and nighttime, representing the typical seasonal variation under undisturbed conditions of the VLF signal amplitude for each considered link. The developed quiet time lines may serve as a tool to determine perturbations of the VLF signal amplitude with solar and geomagnetic as well as atmospheric origin. Also, they allow comparison of the VLF signal amplitude variation for different transmitter-receiver links.

1. Introduction

The D-region, ranging from 50–60 up to 100 km, is the bottom part of the ionosphere. It is weakly ionized and due to its low altitude ranges it has a high density, which causes a large collision frequency of electrons with the main neutral components of the atmosphere. Thus, the D-region is mostly dominated by neutral gas dynamics and chemistry (Baumjohann & Treumann, 1996). During daytime and nighttime, different processes dominate, responsible for the generation of the D-region. This leads to an altitude difference of the D-region ranges between day from 50–60 up to 90–95 km (Pavlov, 2014) and night from ∼80 up to 90–95 km (Thomson et al., 2021). Besides the regular diurnal variation and seasonal differences of the D-region, it also shows ionospheric perturbations with geomagnetic and solar as well as atmospheric origin. Those perturbations of the D-region are also notable in the signal of propagating radio waves with a frequency band of very low frequencies (VLF, 3–30 kHz). Those radio waves are able to travel long distances within the so-called waveguide, confined between Earth's surface and D-region, without much attenuation of only ∼3 dB/1,000 km (Sasmal et al., 2017). VLF waves can have their origin in natural sources, for example, lightning discharges (Ratcliffe, 1966), but they are also generated by man-made transmitters used for naval communication. Particularly single-frequency VLF waves from ground-based and man-made transmitters have a good geographical and often continuous temporal coverage (Thomson et al., 2021). There exist several networks providing long-term VLF measurements (e.g., AARDDVARK, SAVNET and GIFDS).

As a consequence of the above-mentioned D-region altitude variation between day and night, the effective reflection height of the VLF signal also varies; with lower altitudes during daytime (∼65–70 km) and higher altitudes during nighttime (∼80–85 km) considering oblique propagation paths. This results in a general diurnal variation of the VLF signal amplitude, whose signature depends on the east-west components of the transmission path (Clilverd et al., 1999; Ries, 1967). Furthermore, the VLF signal amplitude shows a general seasonal variation, corresponding to the seasonal variation of solar radiation and atmospheric dynamics (Macotela et al., 2021). The signature of the seasonal VLF signal amplitude variation varies between the different transmitter and receiver links, as it also depends on pathway characteristics (e.g., direction, length,

ground conductivity, Richardson and Cohen (2021)). Since the VLF-signal amplitude and phase change are very sensitive to effective reflection height and the vertical electron density gradient, they provide inherent information about the D-region (Wait & Spies, 1964) and its perturbations. Those perturbations are for example, solar flares (Grubor et al., 2008; Wenzel et al., 2016) or gamma-ray bursts (Mondal et al., 2012) as well as by atmospheric effects like lightening (Gołkowski et al., 2021), planetary waves (Schmitter, 2011) or Sudden Stratospheric Warmings (Pal et al., 2017). Although the VLF signal amplitude contains valuable information about the lower ionosphere, it also contains man-made interferences (Biagi et al., 2011). Those interferences are not of interest and often disturb the continuous analysis of these measurements as they don't provide information about the lower ionosphere.

Hence, in this study, we are proposing a method to process the VLF signal amplitudes for retrieving subionospheric information only and generating level 2 data valuable for many different investigations. One valuable product is a quiet time line for each path, representing the typical seasonal variation of the VLF signal amplitude under undisturbed conditions. That is helpful to extract all disturbances for further impact studies. The quiet time has been used for different impact studies, like from solar flares (Wenzel et al., 2016) and geomagnetic storms (Tatsuta et al., 2015), cyclones (NaitAmor et al., 2018; Pal et al., 2020) as well as Sudden Stratospheric Warmings (Pal et al., 2017). Often a fit or an average of the previous quiet days is used (Clilverd et al., 2006; NaitAmor et al., 2018; Pal et al., 2020; Wenzel et al., 2016). Some studies compared impact days with geomagnetic quiet days (Pal et al., 2017; Tatsuta et al., 2015). This works for short events like flares or cosmic rays. However, studying the impact of long-lasting (over a couple of days or longer) geomagnetic storms and atmospheric events can be difficult. It is especially difficult for events occurring around the seasonal transition period, as the amplitude might change significantly according to typical seasonal variation. In the past, there have been introduced various methods to develop a quiet time line for the VLF signal amplitude (Cresswell-Moorcock et al., 2015; Lotz & Clilverd, 2019; Macotela et al., 2021). However, those quiet time lines are generated either year-wise with the focus to rule out solar and geomagnetic disturbances, so the impact of atmospheric events is not averaged and might weigh too strongly. Or, the development of the quiet time line requires continuous and stable VLF data time series, which limits the number of links.

In this paper, we develop the quiet time lines for both, daytime and nighttime, by a polynomial fit of the diurnal and seasonal VLF-signal amplitude composite. For the generation of the composite, we apply a new approach of data preparation, designated as level 2 data, which enables us to use non-continuous and unstable (to a certain degree) VLF data time series. The level 2 data preparation combines handling technically issued amplitude steps by leveling and systematical outlier filtering. The daytime and nighttime VLF signal amplitude quiet time lines for different transmitter-receiver (Tx-Rx) links are determined with that new level 2 data set and presented. Moreover, the quiet time lines are discussed, considering previous publications.

2. Data

In this paper, we illustrate the data processing on different exemplary VLF links. The receiver sites are part of the Antarctic Arctic Radiation-belt Dynamic Deposition VLF Atmospheric Research Konsortia (AARDDVARK), which provides a network of continuous long-range observations of the lower ionosphere in the polar regions (Clilverd et al., 2009). Figure 1 shows the different paths for the links Keflavik, Iceland (Call sign: NRK)—Ny Ålesund (NYA); NRK—Sodankylä (SOD); Cutler, Maine (NAA)—NYA; NAA—SOD and NAA—StJohns (STJ). The chosen links represent VLF signals propagating between high- and middle latitudes, with different path lengths (referring to the great circle path) and characteristics. As shown in Figure 1 the pathways extend over different ground conditions (ice, water, soil), which also affects the VLF-signal propagation due to different conductivity. For instance, the path of NRK-NYA extends mainly over water, which leads to a low attenuation during propagation (Thomson et al., 2011). In Table 1 path lengths, transmitter operating frequencies, midpoint positions, and availability are collected. All here used receivers from AARDDVARK are capable of recording the amplitude and phase of Minimum Shift Keying (MSK) modulated VLF radio transmission (Clilverd et al., 2009). The data analyzed in this study, are 10 min medians of these amplitude data and are denoted as level 0 data in the following. The parameter of 10-min was chosen as it is aimed for a seasonal quiet time variation of the VLF signal. Thus the resolution shouldn't be too high (e.g., 5 min), as especially solar flare impacts would appear quite strongly. The 10-min median work well in this regard, however other time bins (in the minute range) are possible but do not influence the overall variation too much.

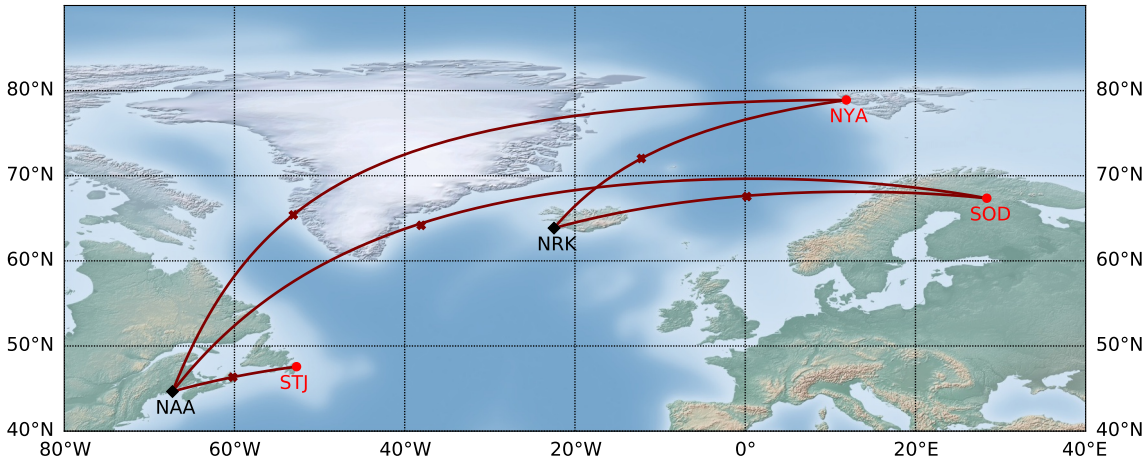

Figure 1. Location of the here considered different Tx-Rx links. Transmitter sites are annotated in black and marked with a diamond, and receiver sites are in red and marked with a dot. The receiver sites belong to the AARDDVARK network. Propagation path midpoints are marked with a maroon cross.

3. Method

This section presents a new approach to processing VLF signal amplitudes. For this, the VLF amplitude jumps in long time series are leveled and the first outliers are filtered with a band-pass filter, resulting in level 1 data. This is followed by a description of the systematic outlier filtering that leads to level 2 data. The approach to defining daytime and nighttime for the seasonal cycle is then described.

3.1. Level 1 Data—Leveling With Change Point Detection Method *PELT*

In long time series, the VLF signal amplitude of various links shows significant jumps. Those jumps, artificially caused by maintenance actions, are manifested by a sudden change of the signal level, while the general signal signature remains unchanged. Especially for the determination of the averaged seasonal behavior of the VLF signal amplitude, but also for the analysis during periods with such a sudden amplitude change, these jumps are problematic. In addition to the jumps, the manifold occurrences of outliers caused by regular and irregular maintenance actions are also a difficulty for analysis. This section uses the VLF signal amplitude of the link NAA-NYA as an example to explain how the amplitude jumps are balanced with the help of the Pruned Exact Linear Timed method (Killick et al., 2012), in the following *PELT*. Furthermore, outlier filtering with a band-pass filter is introduced, which is the first step of a two-step outlier filtering approach. In Figure 2 the blue line represents the level 0 data (10 min medians of raw data amplitude) of the VLF amplitude at 12 LT from 2006 to 2022 for the link NAA-NYA. For example, around mid of the year 2008, there was a sudden decrease, that is, a jump, in the level of the signal notable, whereas the general signature did not change much. Further level jumps occur around 2013 and mid of 2014. In order to balance those jumps, first they need to be identified. The *PELT* method enables to detect jumps in a time series. The basic principle here is to split the regression line of the complete time series into segmented regression lines, whereas the individual regression line segments are determined by minimizing the distances between data points and regression line segments. The so-called change points, which mark the transition between the different regression line segments, thus also represent significant changes in the signal. The

Table 1
Tx-Rx Link Information

Tx-Rx	~path length [km]	Tx frequency [kHz]	Midpoint	Available since [year]
NRK-NYA	2,000	37.5	72.03°N; 12.2°W	2006
NRK-SOD	2,220	37.5	67.55°N; 0.20°E	2010
NAA-NYA	4,930	24	65.4°N; 53.09°W	2006
NAA-SOD	5,660	24	64.17°N; 38.06°W	2010
NAA-STJ	1,170	24	46.34°N; 60.18°W	2013

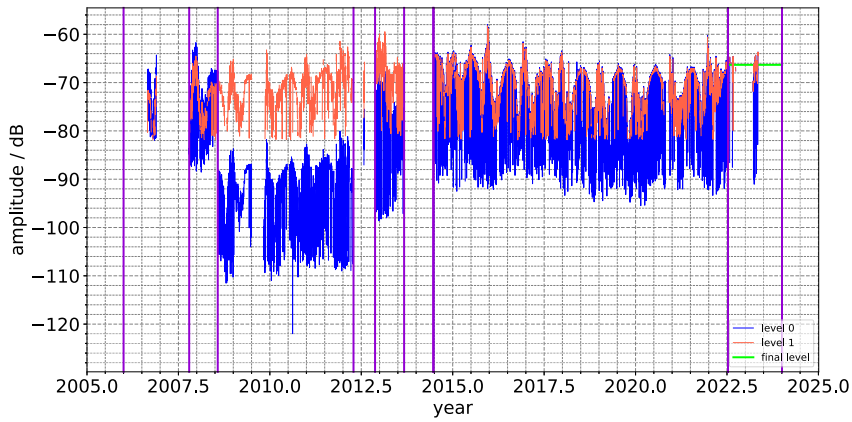

Figure 2. NAA-NYA time series of the VLF signal amplitude at 12 LT. level 0 data (blue) and level 1 data (red), purple vertical lines mark the detected change points. The green line represents the final level on which the complete level 1 time series is set.

PELT method is included in the ruptures python module (Truong et al., 2020). In addition to the time series, the ruptures module also requires a penalty term as an input parameter. The penalty term limits the minimum distance between the change points and so restricts the number of change points. Since the jumps in the data are generated by technical adjustments at either the transmitter or the receiver, the penalty term has to be determined by trial and error for every link individually. For example, it is considered an error if technical jumps are not detected or seasonal amplitude variation is falsely detected as a jump. Generally, the penalty term depends on the number of technically originated jumps, measurement gaps, and the intervals between them. For the link NAA-NYA, the penalty term is 110 time units, here days. The penalty terms for the other links used here in this publication are listed Table S1 in Supporting Information S1.

Since the *PELT* method of the ruptures python module is not able to handle NaNs, the level 0 data gaps are filled with zeros. By setting the missing data to zero they differ strongly from the valid signal values, which supports the change point detection. An interpolation of the data gaps has turned out as not appropriate, especially for small data gaps, since sometimes jumps could not have been identified due to a too-good interpolation. To minimize the risk, that a change point is identified falsely, the *PELT* method is applied for time series at three different times (2, 12, 22 LT) and then compared. If the nighttime and daytime change points do not match, they are not taken into account as it can be assumed that the level jump is not caused by maintenance and should be considered a natural signal variation. The change points detected for the link NAA-NYA are marked with vertical purple lines in Figure 2.

To level the complete time series, it is subdivided into several segments, which are defined by the detected change points. The first and last data points of the time series are set as first and last change point by default. The median of the individual segments is calculated for each time bin. Subsequently, the median over all time bin medians is determined and then subtracted for each segment. By subtracting the median, each segment is leveled. Since some segments contain noise instead of a reliable signal signature, it is not recommended to level every segment. Another parameter is introduced for this purpose, the minimum median length (MML). Once the change points have been detected with the PELT method, a look at the level 0 data segments gives information on whether the data in a segment are true signal measurements or noise. Then the length of the longest noise segment is determined as the minimum median length so that leveling of noise can be prevented. As with the penalty term, the MML for each link has to be determined for every link individually. In further analysis, segments are considered only if their length is larger than the MML, so the short-term noisy segments are ignored. For the link NAA-NYA, the MML is 200 days. As can be seen in Figure 2, there is no critical segment shorter than MML, containing only noise, for the link NAA-NYA. In the time series of the links NAA-SOD (Figure S2 in Supporting Information S1), NRK-SOD (Figure S6 in Supporting Information S1), and NRK-NYA (Figure S4 in Supporting Information S1), the MML was required to prevent the leveling of noise. Both, the time series and the MML (Table S1 in Supporting Information S1) for the other links (Figure S2, Figure S4, Figure S6, and Figure S8) are listed in Supporting Information S1. After the first leveling, the amplitude steps are balanced around the level

zero. Though the jumps are removed, it is difficult to recognize the signal signature, since it is inferior to the many outliers with different origins (also notable in noisy level 0 data, see blue line in Figure 2). It turned out that band-pass filtering (Nisbet et al., 2018) is an effective way to remove outliers, mainly generated by regular and irregular maintenance. Since amplitude values of technically issued outliers and amplitude values during seasonal transition (especially during spring and autumn) can differ only slightly, the particular difficulty here is to choose an appropriate limit, that filters mainly technically issued outliers but not significant data changes, which may occur due to seasonal variation. The limit is determined by the product of the standard deviation of the complete time series and a factor, which depends on the magnitude of the standard deviation by an exponential function. More detailed information on the determination of the factor is described in Supporting Information S1 (Figure S1). Note, that a high standard deviation of the time series does not necessarily mean, that there is a high amount of outliers. A high standard deviation can also be caused by strong but typical seasonal variation. The band-pass filter limit for the link NAA-NYA is ±15.5 dB.

The complete time series is finally set to a level determined by the median of the level 0 data of the last segment (the green line in Figure 2). Here the last segment is chosen as these are the latest data. Note, that the AARDDVARK VLF signal amplitudes are uncalibrated and so a final level is arbitrary, but does provide a framework to compare the relative "day" and "night" levels on each individual path that are shown later on.

The red line in Figure 2 represents the leveled time series of the VLF amplitude at 12 LT (denoted as level 1) for the link NAA-NYA. Compared to the level 0 data (blue line) the jumps have been removed and the complete time series is on one general level. Furthermore, the signal appears cleaner, due to the first outlier filtering. The leveling and the band-pass filtering are carried out for every time series for every 10-min time bin resulting in a level 1 data set.

3.2. Level 2 Data—Outlier Detection With *MAD*

Additionally to the outlier caused by maintenance action, there also may occur extraordinary low or strong amplitude values in the VLF signal amplitude, which are generated by for example, solar storms, flares, or technical disturbances. At least the strongest of those outliers also need to be removed as good as possible, since we are aiming the generation of a quiet time line, which represents the typical seasonal variation under undisturbed conditions. The band-pass filtering is not able to remove such outliers sufficiently, since due to the similar amplitude values of technically issued outliers and amplitude values during seasonal transition periods, there is no large tolerance range for the band-pass filter limit. A second outlier filtering is required. The standard outlier filtering using the mean value and the standard deviation is error-prone for the VLF signal amplitude due to the magnitude of the outlier. Hence the outlier detection by a robust score calculated with the median and the median of all absolute deviations from the median, short *MAD*, is applied (Rousseeuw & Hubert, 2011). The *MAD* is a robust measure of scale, given by:

$$\mathrm{MAD} = 1.483 \underset{\mathbf{i=1,\ldots,n}}{\mathrm{median}} |x_i - \underset{\mathbf{j=1,\ldots,n}}{\mathrm{median}} (x_j)| \tag{1}$$

Where 1.483 works as a correction factor to keep the *MAD* unbiased at the normal distribution, as proposed in Rousseeuw and Hubert (2011). In Equation 1 the n represents the number of considered values, the window size. According to the z-score (standardization) which is a conventional method for outlier detection:

$$z_i = (x_i - \bar{x})/s \tag{2}$$

with s as standard deviation, we apply this rule to the median and *MAD* and obtain the robust score rs:

$$rs_i = \left(x_i - \underset{\mathbf{j=1,\ldots,n}}{\mathrm{median}}(x_j)\right)/\mathrm{MAD} \tag{3}$$

If $|rs_i|$ exceeds a defined cutoff (here 2.85), x_i is flagged as an outlier. The cutoff is for all Tx-Rx links the same and was chosen because it meets the requirement of filtering technical outliers and strong solar and geomagnetic events while at the same time not filtering typical but significant diurnal and seasonal amplitude changes (e.g., at sunrise and sunset or in spring and autumn). With a smaller *MAD* cutoff (<2.85) the natural and significant

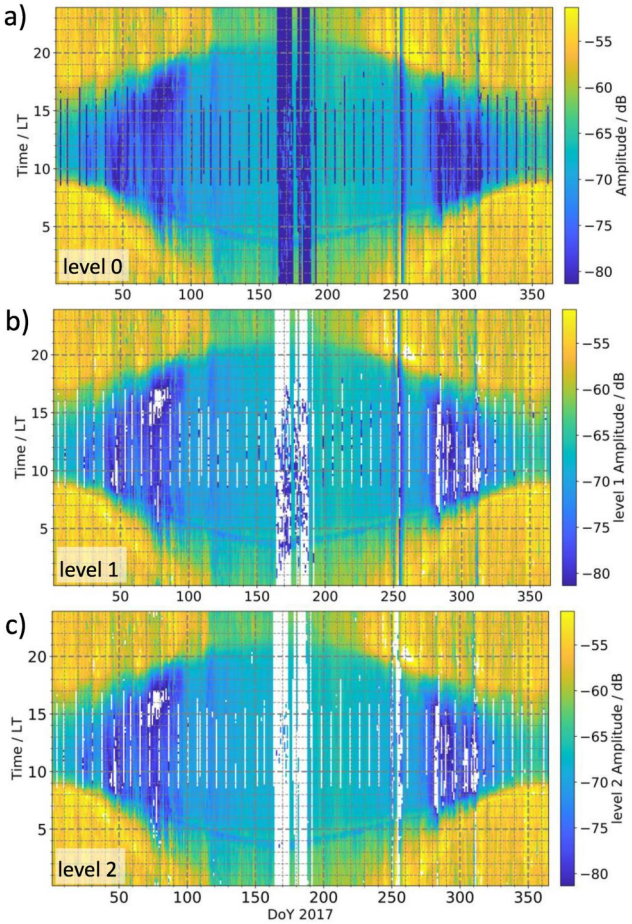

Figure 3. (a) NAA-NYA seasonal and diurnal variation of the VLF amplitude for the year 2017 in three steps: level 0 data, (b) level 1 data with leveling and band-pass filter outlier filtering and (c) level 2 data with an additional outlier filtering using the *MAD* method in two dimensions. Local time (LT) refers to the midpoint position.

seasonal and diurnal variation are falsely identified as an outlier. With a larger *MAD* cutoff (>2.85) not all outliers are detected, especially outliers occurring during longer-lasting (weeks) outages.

To account for seasonal and diurnal variation, the *MAD* method is applied in two dimensions. It is applied first, with a running *MAD* for each daytime separately along the time vector to identify outliers between days. Second, in terms of daytime to identify outliers between years. With a subsequent reconciliation, the outliers are validated, by considering only outliers which match for both methods. This will prevent significant but typical seasonal and diurnal changes from being filtered as outliers (e.g., sunset and sunrise). For the day-wise outlier filtering the robust score is calculated for each time bin in one day steps with a moving *MAD* over a time window of 60 days. A window size of 60 days appears to be an appropriate value for outlier filtering since it removes extraordinarily high or low values but also identifies sunrise and sunset as valid values. With a window size smaller than 60 days, the natural fluctuations between the days would detected falsely as outliers. Whereas in the case of a window size larger than 60 days sunrise and sunset were falsely identified as outliers. For the daytime-wise outlier filtering the *MAD* is calculated for each time bin respectively by considering the median over all available years. According to Equation 1, in the daytime-wise outlier filtering the n is equal the number of years and varies with the data coverage of the considered Tx-Rx link.

In Figure 3 the two-step outlier filtering is presented for the diurnal-seasonal variation of the VLF signal amplitude for NAA-NYA. Figure 3a shows the level 0 data, Figure 3b the level 1 data, which have been already leveled and filtered by the band-pass filter, and Figure 3c shows the level 2 data, where the 2 dimensional *MAD* method have been applied on the level 1 data. In Figure 3a the vertical dark blue lines represent outliers due to regular maintenance (approx. weekly frequency from 9/10–15/16 LT) as well as irregular maintenance work lasting over multiple days (between period DOY: 160–190, and around DOY: 255). Most of those outliers could already be removed by the band-pass filter obtaining level 1 data in Figure 3b. However, level 1 data still show some remaining outliers, which couldn't be filtered sufficiently by the band-pass filter. Most of the remaining outliers are then removed by the *MAD* method, that is, level 2 data. Nevertheless, it must be considered that not all outliers are filtered out completely, especially outliers caused by weaker solar and geomagnetic storms or when they appear around seasonal transition periods due to overlapping effects. This circumstance is considered in the further quiet time generation procedure.

3.3. Definition Daytime and Nighttime

VLF signal amplitude strongly depends on effective reflection height and electron density. Both parameters change significantly with daytime. Hence the VLF signal amplitudes during day and night need to be considered exclusively. Inspired by Politis et al. (2022) we define daytime as the time range when the D-region is sunlit and nighttime as the time range when the D-region is not sunlit. To determine periods, when the altitude ranges of VLF effective reflection heights are sunlit and not sunlit, day and night masks have been developed with the help of different sun position states defined by the solar zenith angle (SZA), which are included in the pure-Python astronomy package "Skyfield" (Rhodes, 2019). All values between the so-called "dark night" (SZA = 108°, in Figure 4 the cyan line) and "nautical twilight dawn" (SZA = 102°, in Figure 4 the dark blue line) are attributed to the nighttime. Values between "civil twilight" (SZA = 96°, in Figure 4 the yellow line) and "nautical twilight dusk" (SZA = 102°, in Figure 4 the red line) are attributed to daytime. Those sun-state positions have been chosen to differentiate more sharply from the twilight period, where the transition from lower daytime to higher nighttime VLF signal amplitudes can be observed. Note that Figure 4 shows the composite of the diurnal and seasonal variation of the VLF amplitude for the link NAA-NYA. How the composite or quiet time lines are determined is

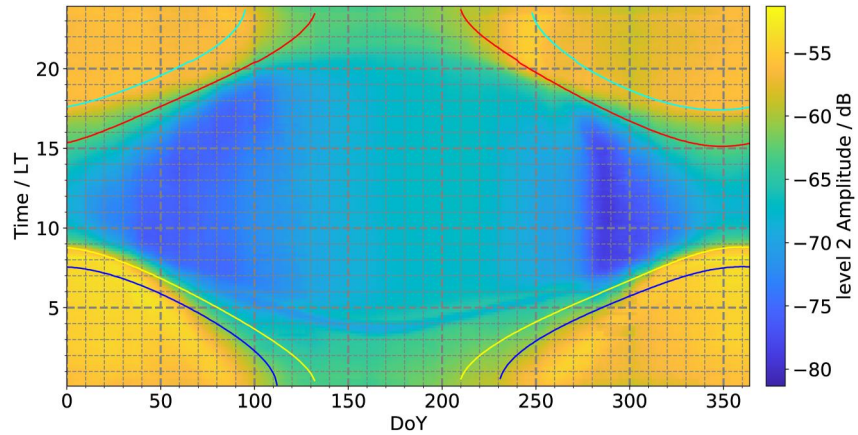

Figure 4. Composite of the diurnal and seasonal variation of the VLF signal amplitude for the link between NAA-NYA. The black boxes mark the time windows used for daytime and nighttime line computation. Colored lines represent different states of SZA: cyan for dark night (SZA = 108°), dark blue for nautical twilight dawn (SZA = 102°), yellow for civil twilight (SZA = 96°), and red for nautical twilight dusk (SZA = 102°). Local time (LT) refers to the midpoint position.

explained in the next section. The temporal development of the different sun state position times for the midpoint position are highlighted as colored lines. The VLF signal amplitude variation matches the determined sun position state times. The transition from lower daytime to higher nighttime VLF signal amplitudes takes place during the twilight periods between nautic twilight dawn (dark blue) and civil twilight (yellow) in the morning and between nautic twilight dusk (red) and dark night (cyan) in the evening.

4. Quiet Day- and Nighttime Seasonal Variation

For the generation of a quiet time line, representing the typical seasonal variation of the VLF signal amplitude under undisturbed conditions, the approach of a superposed epoch analysis is followed. For this purpose, all available level 2 data are stacked yearwise and the median is computed over all years. Hence the composite of the diurnal-seasonal variation is obtained (see Figure 4). The composites for the other links considered here in this study are included in Supporting Information S1 (NAA-SOD in Figure S3, NRK-NYA in Figure S5, NRK-SOD in Figure S7, and NAA-STJ in Figure S9). In Figure 5a the quiet time daytime and nighttime lines for NAA-NYA are depicted, where the quiet time daytime line was computed by the daily median within the range 11–13 LT of the composite data set. The nighttime line was computed by the daily median within the range 22–24 LT of the composite data set. Note that only these values were used which were marked as daytime or nighttime respectively (cf. Section 3). For links in high latitudes, the nighttime line is not continuous as there is a period during summer, where the D-region altitudes don't experience dark night due to continuous sunlight exposure (polar day, cf. Figure 4). The daytime line, on the other hand, is continuous. Even during the polar night period, because the D-region is still illuminated by the sun for a few hours around noon. So daytime effects also occur during the winter time. (cf. Figure 4).

There are not for all Tx-Rx links necessarily sufficient years available, for a statistically robust average representing a quiet time line. Thus single strong events might have a big weight in the average. To smooth the effects of such strong events, the Savitzky-Golay filter (Luo et al., 2005) with a polynomial order 3 and a window size of 17 days is applied to the daytime and nighttime line. A window size of 17 days was determined as appropriate, as it smoothes variations caused by extreme events sufficiently and retains the significant seasonal characteristics in springtime and fall. The smoothed daytime and nighttime lines represent finally the quiet time lines for day and night (red and blue lines in Figure 5a).

In Figure 5 daytime and nighttime quiet time lines for five different links are depicted. The propagation path midpoints of all links lie in polar and mid-latitudes (cf. Figure 1). Figure 5a shows the quiet time line for day- (red) and nighttime (blue) for the link NAA-NYA. The daytime line shows an annual variation with increased amplitudes during summer and winter time, with maximum peaks of −67 dB at 210 DOY and −68 dB at 355 DOY. Compared to summer and winter time the amplitude values in spring and autumn are lower, with negative peaks of −76 dB at 59 DOY and −79 dB at 287 DOY. The amplitude decrease in spring is slower and its magnitude is

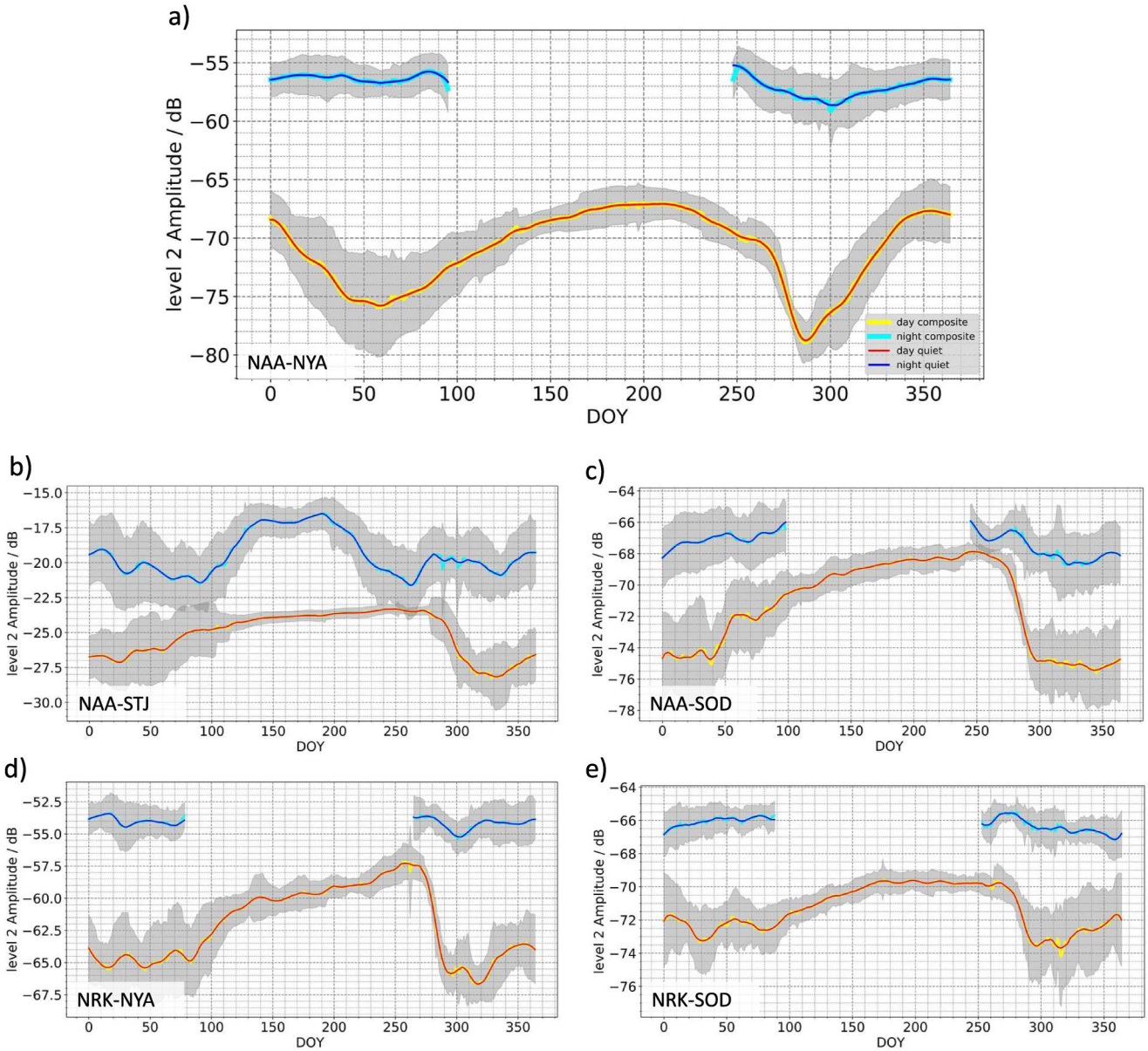

Figure 5. Daytime (yellow) and nighttime (cyan) lines were computed with the help of the composite for the links NAA-NYA (a), NAA-STJ (b), NAA-SOD (c), NRK-NYA (d) and NRK-SOD (e). The respective quiet time lines for daytime (red) and nighttime (dark blue) are generated by an additional smoothing with the Savitzky-Golay filter.

lower than in autumn. In the period 30 days before the minimum peaks until the minimum peaks, the decrease in spring is approximately −0.1 dB/day, whereas the decrease in autumn is approximately −3 dB/day. This strong decrease in autumn is known as the October effect (Pancheva & Mukhtarov, 1996). The asymmetry between spring and autumn has also been observed in Macotela et al. (2021). The standard deviation for the daytime line is higher in winter (approx. 2–4 dB) than in summer (approx. 1–2 dB), which is an indicator of an atmospheric dynamic impact on the VLF signal amplitude. During wintertime, the planetary wave activity is increased, due to changed filter conditions for planetary wave propagation (Charney & Drazin, 1961). According to Schmitter (2011) the increased wave activity correlates with the often described winter anomaly. Note, that the nighttime line is not continuous during the year, since there is polar day in summer at the midpoint position, which is located in high latitudes. For winter, spring, and autumn the nighttime line does not show strong variations. Only a slight decrease in the amplitude is observable during autumn. The nighttime line shows higher amplitude values than the

daytime line, which is also observed in Clilverd et al. (2006) and explained by a stronger attenuation of the signal on the propagation path during daytime due to photo-ionization.

For all links in Figure 5 the daytime quiet time lines show the typical seasonal variation which in general follows the solar zenith angle, with increased amplitudes in the summertime. The peculiarity of the spring-fall asymmetry (Macotela et al., 2021) is also notable in all daytime quiet time lines, whereas the strength of the asymmetry varies between different links. For all links, the standard deviation, represented by the shaded areas in Figure 5, is higher in winter than in summer which likely has its origin in the winter anomaly. Furthermore, there are higher nighttime amplitudes than daytime amplitudes for all links. The daytime quiet time lines in Figures 5b–5e differ from the NAA-NYA daytime line in Figure 5a as their amplitudes have only one maximum in summer, whereas the NAA-NYA daytime line shows a semiannual behavior with a maximum in summer and in winter. The daytime quiet time lines for the links NRK-NYA and NRK-SOD (see Figure 5 d,e) indicate a pronounced wave structure during wintertime. Those wave structures are also notable, though weaker, in the daytime quiet time lines for NAA-STJ and NAA-SOD (see Figures 5b and 5c). The nighttime quiet time line for NAA-STJ also shows a pronounced wave structure. It seems that the wave structure has its maximum in higher altitudes since it is stronger in the nighttime signal of NAA-STJ and in the daytime lines for NRK-NYA and NRK-SOD. The transmitter NRK operates with a frequency of 37.5 kHz, which is higher than NAA with 24 kHz. Accordingly, the daytime effective reflection height is higher for the VLF signal for the transmitter NRK. Except for the nighttime quiet time line for the link NAA-STJ, all other nighttime lines are not continuous due to polar day. Note, that the continuous nighttime quiet time line for NAA-STJ does not show the spring-autumn asymmetry, which is observed for all links during daytime. This suggests that the spring-autumn asymmetry occurs only in the lower daytime effective reflection heights.

In this section, quiet time lines for the day- and nighttime for different Tx-Rx links were developed with the help of the level 2 data set. The quiet time lines show typical seasonal variation with at least one VLF amplitude maximum in summer and a spring-fall asymmetry during daytime.

5. Discussion and Concluding Remarks

In this study VLF amplitude measurements for five different links have been used to develop daytime and nighttime quiet time lines for each link respectively, representing the typical seasonal variation of the VLF amplitude under undisturbed conditions. In order to develop the quiet time line by a polynomial fit of the composite, the VLF amplitude data needed to be processed, to balance out technically issued amplitude jumps as well as to filter sufficiently the outlier. With the help of the *PELT* method, the signal was leveled. Subsequently, a band-pass filter and the *MAD* method were applied in two dimensions, to systematically filter the outlier. To all links the same outlier filtering was applied, to obtain comparable results. Even if the outlier filtering has the focus on artificially caused outlier, the VLF signal amplitude variation due to extraordinarily strong solar and geomagnetic events might also be recognized as an outlier in rare cases. For analyzing the effects of extraordinarily strong solar and geomagnetic events, the level 1 data may be better suited, as only band-pass filtering was applied here to eliminate mainly technically induced outliers. If the application of the level 2 data requires stricter filtering of outliers with space weather origin, a filtration based on the time series of space weather events, for example, the x-ray flux or energetic particle precipitation, might be a valuable extension. Though the *PELT* method turned out to be suitable for identifying significant changes in the VLF amplitude, the necessary parameters need to be chosen carefully to prevent leveling of noise as well as too strong leveling. Since the amplitude jumps have a technical origin, they differ between different transmitter and receiver combinations. Hence the *PELT* parameters must be determined for each link respectively. This is time-consuming at first but pays off afterward as the *PELT* method can be applied to the complete time series covering several years. Assuming the most appropriate *PELT* parameter has been determined, it is still uncertain if the median of the last level segment, which is used as a final level for the complete time series, represents the real VLF signal amplitude level, due to missing quality parameter of the VLF measurements. It could just be assumed that the latest longer-lasting segment of VLF amplitude measurements is the most reliable, due to the latest technical improvements. So, changes in the amplitude can be compared for different links, whereas a comparison of absolute amplitude values for different links requires additional consideration of the level 0 data. Even if the pre-processing of the data enables the computation of the quiet time lines with composites for various VLF links, it cannot apply successfully to all data sets. The data need to fulfill some minimum requirements. The general

signature of the signal should not change significantly, which can be caused for instance by changing the sound card. Furthermore, data coverage over at least 2 years is required. However, from a statistical point of view, the principle applies: the more years, the better. The new approach for VLF signal amplitude processing is presented here for data based on AARDDVARK observations only. It is assumed that data based on other observations, for example, AWESOME, SAVNET, or GIFDS, can also be processed with this approach, as long the data coverage is sufficient. However, the specific penalty term and minimum median length for the *PELT* method need to be determined for the new data sets. Furthermore, the parameters for the outlier filtering need to be checked and adjusted if necessary. There are only a few publications about long-term quiet time lines representing the diurnal and seasonal variation. For instance, Cresswell-Moorcock et al. (2015) compared three different techniques to determine the quiet day curve for a long period of subionospheric VLF observations and achieved good results for the combination of a pre-smoothing process and a fast Fourier transform-based quiet day curve finding technique. This approach is especially suitable in the case of a data set without amplitude jumps and considering VLF perturbations for a specific year. For a comparison of the VLF amplitude variation between different years, an averaged quiet time line would be beneficial, since in a year-wise quiet time line the impact of strong atmospheric events such as Sudden Stratospheric Warmings cannot be distinguished. Atmospheric events are able to have an impact with magnitudes comparable to geomagnetic storms (Goncharenko et al., 2010). Thus in generating a quiet time line, additionally to the solar and geomagnetic conditions, the atmospheric conditions should also be taken into account. Macotela et al. (2021) has considered solar and geomagnetic as well as atmospheric conditions by computing the quiet time line by the average over all available years. This approach is very similar to ours, apart from the preliminary data processing with leveling and the outlier filtering, introduced in this study. In fact, the created daytime quiet time line for the link NAA-SOD in this paper (cf. Figure 5b) is similar to the quiet time line in Macotela et al. (2021) (their Figure 1b), where they used only links without jumps. This indicates that the systematically preliminary data processing is a valid method, which enables composite studies and allows comparison between different links.

With the data processing using *PELT* to level the data (identify jumps), a band-pass filter, and the *MAD* method to filter outliers, we developed an approach to use data sets over longer time periods even with amplitude jumps. This enables in general a systematic analysis of long-term VLF observations. Here we applied the composite analysis to get the seasonal variation of the VLF amplitude. The here-developed seasonal variation represents a helpful tool for identifying perturbations with solar- and geomagnetic as well as atmospheric origin. Moreover, the developed quiet time lines deliver information about the annual VLF amplitude signature for each link respectively. Furthermore, the explanation of the differences between the quiet time lines themselves, would also improve the general understanding of the drivers of VLF signal amplitude variation, but goes beyond the scope of this paper. It must be taken into account that the quiet time lines were created using a composite and therefore the solar cycle is not depicted. However, a comparison between quiet time lines and VLF amplitudes under consideration of the state of the solar cycle could also provide useful information to study the solar cycle dependency of the VLF signal amplitude. Such a study is exemplified for the link NAA-SOD in Neal et al. (2015) and indicates a solar cycle dependency. Since the same approach for data processing is applied, the seasonal variations are comparable in a matter of characteristic. The analysis of the similarities and differences between different links may also be useful to gain a further understanding of the interactions between D-region dynamics, ionization, path characteristics, and the VLF signal amplitude.

Data Availability Statement

VLF signal amplitude data from the AARDDVARK network were downloaded from https://psddb.nerc-bas.ac.uk. The link leads to a landing page with a search mask in the upper right corner. The receiver station must be entered into the search mask to access the data. Here, choosing the year and the data class *ULTRA* leads to an information page from where you can access data to all transmitter variations to the chosen receiver for the determined year under the term *View this Data*. For example, here the direct link to the Ny Ålesund data: https://psddb.nerc-bas.ac.uk/data/access/download.php?searchterm=Ny%20Alesund&page=1&cat=item&year=2007&class=232&type=ULTRA&site=Ny-Alesund&v=Data.

Acknowledgments
We thank David Wenzel and Marc Hansen for the fruitful discussions. We also thank all ARRDDVARK members of the Konsortia and the University of Oulu (operating through the Sodankylä Geophysical Observatory) for maintaining the VLF receiver and the provision of data. Open Access funding enabled and organized by Projekt DEAL.

References

Baumjohann, W., & Treumann, R. A. (1996). *Basic space plasma physics* (pp. 74–75). Imperial College Press.
Biagi, P., Maggipinto, T., Righetti, F., Loiacono, D., Schiavulli, L., Ligonzo, T., et al. (2011). The European VLF/LF radio network to search for earthquake precursors: Setting up and natural/man-made disturbances. *Natural Hazards and Earth System Sciences*, *11*(2), 333–341. https://doi.org/10.5194/nhess-11-333-2011
Charney, J. G., & Drazin, P. G. (1961). Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, *66*(1), 83–109. https://doi.org/10.1029/JZ066i001p00083
Clilverd, M. A., Rodger, C. J., Thomson, N. R., Brundell, J. B., Ulich, T., Lichtenberger, J., et al. (2009). Remote sensing space weather events: Antarctic-arctic radiation-belt (dynamic) deposition-VLF atmospheric research konsortium network. *Space Weather-The International Journal of Reasearch and Applications*, *7*(4), S04001. https://doi.org/10.1029/2008SW000412
Clilverd, M. A., Seppälä, A., Rodger, C. J., Verronen, P. T., & Thomson, N. R. (2006). Ionospheric evidence of thermosphere-to-stratosphere descent of polar NOX. *Geophysical Research Letters*, *33*(19), L19811. https://doi.org/10.1029/2006GL026727
Clilverd, M. A., Thomson, N. R., & Rodger, C. J. (1999). Sunrise effects on VLF signals propagating over a long north-south path. *Radio Science*, *34*(4), 939–948. https://doi.org/10.1029/1999RS900052
Cresswell-Moorcock, K., Rodger, C. J., Clilverd, M. A., & Milling, D. (2015). Techniques to determine the quiet day curve for a long-period of subionospheric VLF observations. *Radio Science*, *50*(5), 453–468. https://doi.org/10.1002/2015RS005652
Gołkowski, M., Renick, C., & Cohen, M. B. (2021). Quantification of ionospheric perturbations from lightning using overlapping paths of VLF signal propagation. *Journal of Geophysical Research: Space Physics*, *126*(5), e2020JA028540. https://doi.org/10.1029/2020JA028540
Goncharenko, L. P., Coster, A. J., Chau, J. L., & Valladares, C. E. (2010). Impact of sudden stratospheric warmings on equatorial ionization anomaly. *Journal of Geophysical Research*, *115*(A10). https://doi.org/10.1029/2010JA015400
Grubor, D. P., Šulić, D. M., & Žigman, V. (2008). Classification of x-ray solar flares regarding their effects on the lower ionosphere electron density profile. *Annales Geophysicae*, *26*(7), 1731–1740. https://doi.org/10.5194/angeo-26-1731-2008
Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.48550/arXiv.1101.1438
Lotz, S. I., & Clilverd, M. (2019). Demonstrating the use of a class of min-max smoothers for d region event detection in narrow band VLF phase. *Radio Science*, *54*(3), 233–244. https://doi.org/10.1029/2018RS006701
Luo, J., Ying, K., & Bai, J. (2005). Savitzky–golay smoothing and differentiation filter for even number data. *Signal Processing*, *85*(7), 1429–1434. https://doi.org/10.1016/j.sigpro.2005.02.002
Macotela, E. L., Clilverd, M., Renkwitz, T., Chau, J., Manninen, J., & Banys, D. (2021). Spring-fall asymmetry in VLF amplitudes recorded in the north Atlantic region: The fall-effect. *Geophysical Research Letters*, *48*(16). https://doi.org/10.1029/2021GL094581
Mondal, S. K., Chakrabarti, S. K., & Sasmal, S. (2012). Detection of ionospheric perturbation due to a soft gamma ray repeater SGR J1550-5418 by very low frequency radio waves. *Astrophysics and Space Science*, *341*(2), 259–264. https://doi.org/10.1007/s10509-012-1131-5
NaitAmor, S., Cohen, M. B., Kumar, S., Chanrion, O., & Neubert, T. (2018). VLF signal anomalies during cyclone activity in the Atlantic Ocean. *Geophysical Research Letters*, *45*(19). https://doi.org/10.1029/2018GL078988
Neal, J. J., Rodger, C. J., Clilverd, M. A., Thomson, N. R., Raita, T., & Ulich, T. (2015). Long-term determination of energetic electron precipitation into the atmosphere from AARDDVARK subionospheric VLF observations. *Journal of Geophysical Research: Space Physics*, *120*(3), 2194–2211. https://doi.org/10.1002/2014JA020689
Nisbet, R., Miner, G., & Yale, K. (2018). *Handbook of statistical analysis and data mining applications* (2nd ed., pp. 55–82). Academic Press. https://doi.org/10.1016/B978-0-12-416632-5.00004-9
Pal, S., Hobara, Y., Chakrabarti, S. K., & Schnoor, P. W. (2017). Effects of the major sudden stratospheric warming event of 2009 on the subionospheric very low frequency/low frequency radio signals. *Journal of Geophysical Research: Space Physics*, *122*(7), 7555–7566. https://doi.org/10.1002/2016JA023813
Pal, S., Sarkar, S., Midya, S. K., Mondal, S. K., & Hobara, Y. (2020). Low-latitude VLF radio signal disturbances due to the extremely severe cyclone fani of May 2019 and associated mesospheric response. *Journal of Geophysical Research: Space Physics*, *125*(5). https://doi.org/10.1029/2019JA027288
Pancheva, D., & Mukhtarov, P. Y. (1996). Modelling of the electron density height profiles in the mid-latitude ionospheric d-region. *Annales Geophysicae*, *39*(4). https://doi.org/10.4401/ag-4021
Pavlov, A. (2014). Photochemistry of ions at d-region altitudes of the ionosphere: A review. *Surveys in Geophysics*, *35*(2), 259–334. https://doi.org/10.1007/s10712-013-9253-z
Politis, D. Z., Potirakis, S. M., Sasmal, S., & Hayakawa, M. (2022). Long term variation of VLF signal characteristics during sunrise and sunset times as observe from Athens (GREECE). In *Presentation at COSPAR*.
Ratcliffe, J. A. (1966). Whistlers and related ionospheric phenomena. *Geophysical Journal International*, *11*(5), 563–564. https://doi.org/10.1111/j.1365-246X.1966.tb03172.x
Rhodes, B. (2019). *Skyfield: High precision research-grade positions for Planets and Earth satellites generator*. Astrophysics Source Code Library.
Richardson, D. K., & Cohen, M. B. (2021). Seasonal variation of the d-region ionosphere: Very low frequency (VLF) and machine learning models. *Journal of Geophysical Research: Space Physics*, *126*(9). https://doi.org/10.1029/2021JA029689
Ries, G. (1967). Results concerning the sunrise effect of VLF signals propagated over long paths. *Radio Science*, *2*(6), 531–538. https://doi.org/10.1002/rds196726531
Rousseeuw, P., & Hubert, M. (2011). Robust statistics for outlier detection. *WIREs Data Mining and Knowledge Discovery*, *1*(1), 73–79. https://doi.org/10.1002/widm.2
Sasmal, S., Basak, T., Chakraborty, S., Palit, S., & Chakrabarti, S. K. (2017). Modeling of temporal variation of very low frequency radio waves over long paths as observed from Indian Antarctic stations. *Journal of Geophysical Research: Space Physics*, *122*(7), 7698–7712. https://doi.org/10.1002/2016JA023812
Schmitter, E. D. (2011). Remote sensing planetary waves in the midlatitude mesosphere using low frequency transmitter signals. *Annales Geophysicae*, *29*(7), 1287–1293. https://doi.org/10.5194/angeo-29-1287-2011
Tatsuta, K., Hobara, Y., Pal, S., & Balikhin, M. (2015). Sub-ionospheric VLF signal anomaly due to geomagnetic storms: A statistical study. *Annales Geophysicae*, *33*(11), 1457–1467. https://doi.org/10.5194/angeo-33-1457-2015
Thomson, N. R., Clilverd, M., Brundell, J. B., & Rodger, C. (2021). Quiet night arctic ionospheric d region characteristics. *Journal of Geophysical Research: Space Physics*, *126*(4). https://doi.org/10.1029/2020JA029043

Thomson, N. R., Rodger, C. J., & Clilverd, M. A. (2011). Daytime d region parameters from long-path VLF phase and amplitude. *Journal of Geophysical Research*, *116*(A11). https://doi.org/10.1029/2011JA016910

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. https://doi.org/10.1016/j.sigpro.2019.107299

Wait, J. R., & Spies, K. P. (1964). *Characteristics of the Earth-ionosphere waveguide for VLF radio waves*. National Bureau of Standards.

Wenzel, D., Jakowski, N., Berdermann, J., Mayer, C., Valladares, C., & Heber, B. (2016). Global ionospheric flare detection system (GIFDS). *Journal of Atmospheric and Solar-Terrestrial Physics*, *138–139*, 233–242. https://doi.org/10.1016/j.jastp.2015.12.011