UNIVERSITY OF BIRMINGHAM

Wavelet Analysis of Differential TEC Measurements Obtained Using LOFAR

Boyde, Ben; Wood, Alan; Dorrian, Gareth; Sweijen, Frits; de Gasperin, Francesco; Mevius, Maaijke; Beser, Kasia; Themens, David

Document Version
Publisher's PDF, also known as Version of record

[Link to publication on Research at Birmingham portal](#)

Wavelet Analysis of Differential TEC Measurements Obtained Using LOFAR

Ben Boyde[1] , Alan Wood[1] , Gareth Dorrian[1] , Frits Sweijen[2], Francesco de Gasperin[3,4], Maaijke Mevius[5] , Kasia Beser[5], and David Themens[1]

[1]Space Environment and Radio Engineering, School of Engineering, University of Birmingham, Birmingham, UK, [2]Leiden Observatory, Leiden University, RA Leiden, The Netherlands, [3]INAF - Istituto di Radioastronomia, Bologna, Italy, [4]Hamburger Sternwarte, University of Hamburg, Hamburg, Germany, [5]ASTRON – The Netherlands Institute for Radio Astronomy, PD Dwingeloo, The Netherlands

Abstract Radio interferometers used to make astronomical observations, such as the LOw Frequency ARray (LOFAR), experience distortions imposed upon the received signal due to the ionosphere as well as those from instrumental errors. Calibration using a well-characterized radio source can be used to mitigate these effects and produce more accurate images of astronomical sources, and the calibration process provides measurements of ionospheric conditions over a wide range of length scales. The basic ionospheric measurement this provides is differential Total Electron Content (TEC, the integral of electron density along the line of sight). Differential TEC measurements made using LOFAR have a precision of <1 mTECu and therefore enable investigation of ionospheric disturbances which may be undetectable to many other methods. We demonstrate an approach to identify ionospheric waves from these data using a wavelet transform and a simple plane wave model. The noise spectra are robustly characterized to provide uncertainty estimates for the fitted parameters. An example is shown in which this method identifies a wave with an amplitude an order of magnitude below those reported using Global Navigation Systems Satellite TEC measurements. Artificially generated data are used to test the accuracy of the method and establish the range of wavelengths which can be detected using this method with LOFAR data. This technique will enable the use of a large and mostly unexplored data set to study traveling ionospheric disturbances over Europe.

1. Introduction

The ionosphere contains variations on a broad range of length scales, from 10s of meters to 1000s of kilometres. These perturbations in plasma density lead to perturbations in refractive index at radio frequencies, meaning that radio signals passing through the ionosphere will be distorted by these ionospheric structures (Aarons, 1982). To first order, these distortions are phase changes proportional to the integrated electron density along the line of sight, known as Total Electron Content (TEC). Higher order terms include magnetic field effects and non-linear integrals of electron density. All these ionospheric effects have characteristic frequency dependencies, which can be used to separate ionospheric distortions from distortions arising from instrumental effects (de Gasperin et al., 2019, and references therein).

Ionospheric disturbances include both turbulent structures associated with various plasma instability processes, and coherent wave structures. These waves are known as traveling ionospheric disturbances (TIDs) and are typically generated by atmospheric gravity waves (AGWs) propagating in the neutral atmosphere (Hines, 1960; Hocke & Schlegel, 1996). AGWs in the thermosphere have a range of sources, including convective weather systems in the troposphere, winds flowing over topographical features, and Joule heating at high latitudes due to geomagnetic activity (Fritts & Alexander, 2003). Observations of TIDs are key to understanding the dynamics of the ionosphere-thermosphere system, especially due to the difficulty of directly observing AGWs themselves in the thermosphere.

TIDs are commonly observed using a range of data sources, including ionosondes, incoherent scatter radars and TEC derived from Global Navigation Systems Satellite (GNSS) signals (e.g., Lan et al., 2018; Munro, 1950; Otsuka et al., 2013; Themens et al., 2022; van de Kamp et al., 2014). These have revealed TIDs with typical horizontal wavelengths ranging from ∼100 to 1,000 km and periods ranging from roughly 10 min to 2 hr. These are commonly divided into Medium-Scale TIDs (MSTIDs) with wavelengths of 100−300 km and Large-Scale

TIDs (LSTIDs) with wavelengths above 300 km (Hunsucker, 1982). The lower wavelength limit of 100 km is largely due to the resolution limitations of the most widely available observations.

The LOw Frequency ARray (LOFAR: van Haarlem et al., 2013) is a radio telescope centered in the Netherlands, with international stations spread as far as Ireland and Latvia. The stations within the Netherlands are divided into "core" stations, which all lie within a roughly 3 km radius, and "remote" stations up to ∼60 km from the core. Stations are identified by two letters and three numbers (e.g., CS002, RS306): CS and RS denote core and remote stations respectively while international stations are denoted by a two letter country code. LOw Frequency ARray can observe in three frequency bands, 10 MHz −90 MHz, 110 MHz −190 MHz and 210 −250 MHz. While these baselines ranging from 10s of meters to 100s of kilometres function well for use within an interferometer for astronomical observations, they also mean that ionospheric conditions along the line of sight from different stations can differ significantly.

Although it is primarily intended for astronomical observations (e.g., Carbone et al., 2016; Heald et al., 2015; Shimwell et al., 2017; Yatawatta et al., 2013), the distortions to the observed signals mean that LOFAR can also be used for ionospheric observations. For example, Fallows et al. (2020) were able to achieve the first detection of two simultaneous TIDs at different altitudes and traveling in different directions by analyzing variations in received signal intensity observed using LOFAR. In another observation, Boyde et al. (2022) were able to show that their observed intensity variations corresponded to focusing from a small-scale TID with a wavelength of ∼15−30 km. Another example of small-scale structure within a TID detected using LOFAR is the ∼20 km substructure within an MSTID reported by Dorrian et al. (2023).

In order to mitigate both ionospheric and instrumental distortions to the observed signals, typical astronomical observations with LOFAR include a calibrator source that is observed simultaneously or in series with the target field (de Gasperin et al., 2019). These calibrators are bright, compact, well-characterized sources, meaning that the observed signal can be compared to the expected signal to estimate the complex gain, a process known as self-calibration (Pearson & Readhead, 1984). This complex gain must be separated into contributions from different effects, due to the fact that some of these effects vary depending on the viewing direction. As the calibrator dominates over other sources in the field of view, the direction dependent ionospheric effects can be treated as direction independent in the calibration, as only the contribution toward the calibrator is significant and we can ignore variations across the field of view. Effects that are truly direction independent, such as many instrumental effects, can then be corrected before calibrating the target field.

The LOFAR calibration process typically provides three ionospheric measurements, which are as follows: differential TEC (dTEC, the difference in TEC observed by two different stations), differential Faraday rotation (dFR, the difference in Faraday rotation observed by two different stations) and intensity scintillation (variations in signal intensity due to focusing/defocusing in the ionosphere). In addition to these, the third order ionospheric delay, describing the effects of concentration of ionospheric density in thin layers (Hoque & Jakowski, 2008), is also calculated for observations below 40 MHz (de Gasperin et al., 2018). The phase effects (dTEC and dFR) can only be determined as differential quantities because the "original" signal phase from any natural radio source is unknown but can be assumed to be spatially uniform when incident on the ionosphere over the size of LOFAR. As a result, one station must be arbitrarily defined as having zero phase and then all others can be compared to this. In this paper we focus entirely on the dTEC measurements as these are the most directly applicable to quantifying ionospheric disturbances.

While absolute measurements of TEC such as those derived from dual-frequency GNSS receivers are vital for many applications, the differential measurements provided by LOFAR contain all the necessary information for studying disturbances if knowledge of the background conditions is not required. The quality of dTEC derived from LOFAR observations was studied by Mevius et al. (2016), who found a noise floor of <1 mTECu (1 TECu = 10^{16}e⁻ m⁻²), an order of magnitude better than GNSS derived TEC (Otsuka et al., 2013). The phase residuals after all corrections were applied were found to vary systematically with pointing direction across multiple observations. This suggests that this noise floor is a result of either an incomplete sky model, an inaccurate model of the beam shape or some combination of the two. The structure function of the TEC fluctuations was found to be consistent with a mixture of Kolmogorov turbulence and coherent wave activity, and typically anisotropic, aligning with the geomagnetic field more often than not (Mevius et al., 2016).

dTEC measurements from radio interferometers have been used to study waves in the ionosphere in the past. For example, Jacobson and Erickson (1992) used Fourier analysis on data from the Very Large Array (VLA: Thompson et al., 1980) to identify a range of waves, including fast disturbances which appeared to propagate toward magnetic east and were later shown to be in the plasmasphere (Hoogeveen & Jacobson, 1997; Jacobson & Erickson, 1993). More recently, their approach was refined by Helmboldt et al. (2012) to relax some of the constraints that had been imposed on possible wave solutions. Using LOFAR, Beser et al. (2022) were able to identify variations in dominant wave direction over the course of several hours and show that these tracked variations in the geomagnetic field and plasma convection.

This paper presents a novel analysis method for identifying the parameters of ionospheric waves from dTEC measured using LOFAR. This includes robust estimation of the uncertainty in the derived wave parameters. An example data set is shown in which this approach identifies a wave with amplitude smaller than the sensitivity of relative TEC from GNSS measurements (e.g., Otsuka et al., 2013). The paper is structured as follows: first, the observations used are introduced and the characteristics of waves in dTEC measurements are described. Then the analysis method is introduced, including the procedure for uncertainty estimation, and is demonstrated on the observations. Following this, synthetic data are used to examine the performance of the method in identifying waves of different parameters in LOFAR data. Finally, the prospects for applying this method to a wider data set are discussed.

2. Observations

To illustrate the analysis methods described in this paper, we will use the calibration solution for an 8-hr observation (ID L691726) made using the LOFAR high-band array between 120.2 and 187.4 MHz with a frequency resolution of ~48.8 kHz. This observation ran from 15:15 UT to 23:15 UT on 21 December 2018, and targeted the source 3C48 (RA 1h37m41s, Dec. 33.16°). As 3C48 is one of the regular calibrator sources, dTEC solutions were able to be derived directly for the target source in this case. While the international stations were included in this observation, here we focus only on the stations in the Netherlands where coherent structures are more likely to span all lines of sight. Over the course of the observation, the apparent position of the source moved from due East (~90° azimuth) to due West (~270° azimuth) as viewed from the LOFAR core, starting and ending at an elevation of ~44° and reaching a maximum elevation of 70°.

The dTEC solutions for the full observing window for a range of stations are shown in Figure 1. These show activity across timescales ranging from minutes to hours. The baselines within the core (top panel) are generally all very closely correlated with one another. Some remote station baselines (bottom panel) are also well correlated, such as those to RS310 and RS409, due to their similar orientation and length. However, there are also several features in the remote station baselines that are unique to a single baseline, such as the brief, sharp increase in dTEC on the baseline to RS210 just after 17:00 UT. It is also important to note that the magnitudes of the dTEC values here, even on the longest baselines to the remote stations, are significantly smaller than the magnitude of variations that have been shown elsewhere (see e.g., de Gasperin et al., 2018; Mevius et al., 2016). This is likely due in part to the observation being made at solar minimum during winter solstice and largely at night, meaning ionospheric densities and therefore TEC and TEC variations are particularly small. As a result, these data should provide a good test case to demonstrate the sensitivity of these measurements to even such small variations in TEC.

Within these dTEC solutions, we focus on one specific time period between roughly 17:20–17:50 UT during which there is clear wave activity across the array. The full wave progression is shown in Movie S1 in the Supporting Information, while Figure 2 shows the detrended dTEC values at four times during the passage of this wave starting from the appearance of an enhancement in the westernmost baseline and showing it passing into the eastern baselines. While this wave was observed, the line of sight was to the South-East of the LOFAR core (azimuth ~130°), at an elevation of ~64°. The detrending is carried out by subtracting a 60 min running average from the dTEC value on each baseline, to remove large scale diurnal variations while retaining shorter period wave signals. The wave clearly propagates from west to east and with a period of roughly 10 min, although it should be noted that the period will be Doppler shifted due to the apparent scan velocity of the line of sight through the ionosphere, typically of order ~15−20 m s^{-1} around the F-region electron density peak. However, this visual inspection does not permit determination of the wavelength or amplitude, which require a more systematic approach which will be described in the following sections. The amplitude of the variations measured here is close to or somewhat below the precision of GNSS measurements (Otsuka et al., 2013), demonstrating the value of the high precision measurements provided by LOFAR (Mevius et al., 2016).
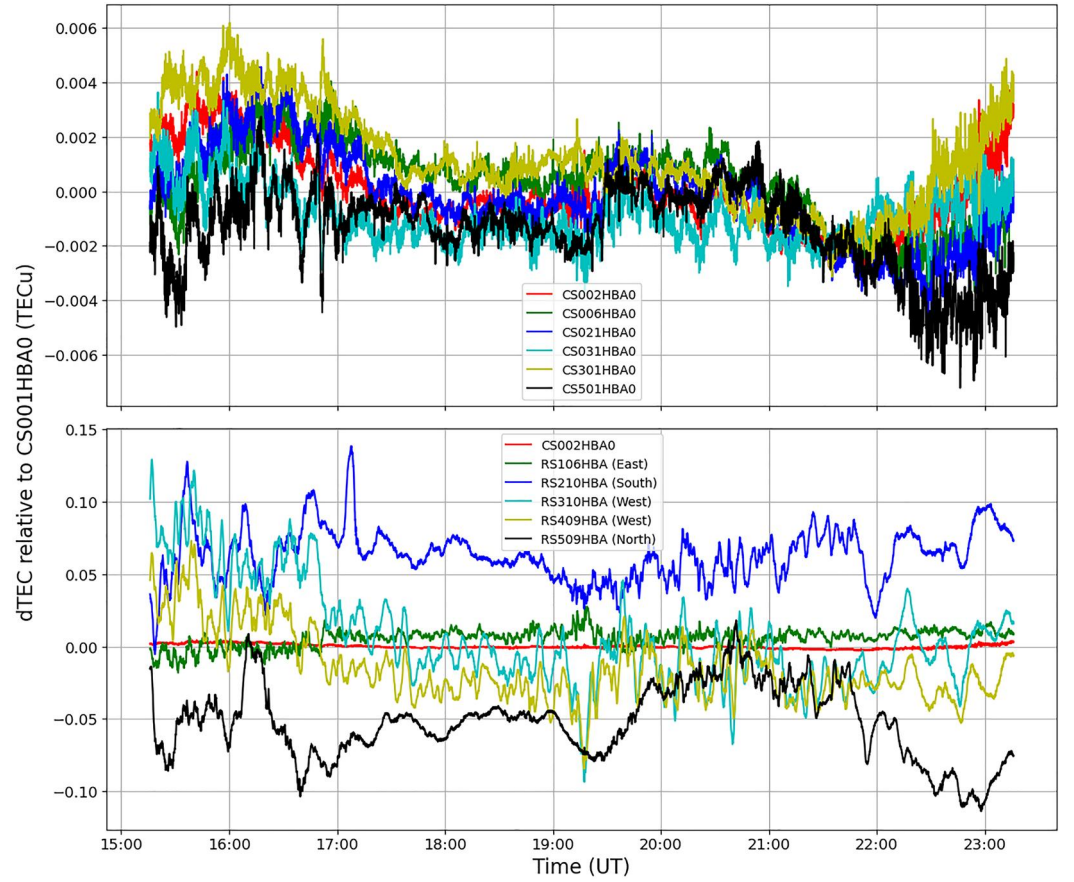
Figure 1. A selection of dTEC time series for the whole of observation L691726 on 21 December 2018, referenced to CS001HBA0. The top panel shows the dTEC on various baselines within the core, while the bottom panel shows dTEC on baselines to remote stations with one core station baseline for comparison. Note the difference in magnitude of the vertical scales in each panel. The remotes shown are the furthest from the core along their respective directions.

Given the combination of high measurement precision and high density of measurements in the Netherlands provided by LOFAR, the dTEC data provide an opportunity to extend the parameter space in which we can observe ionospheric waves. However, the use of differential rather than absolute TEC measurements presents some additional challenges in analysis. These differences impose certain limitations on the parameter space that can be accessed by LOFAR which do not apply to GNSS measurements. The difference between the wave signal seen in GNSS absolute TEC and differential TEC is shown in Figure 3 for an assumed plane wave. The absolute TEC signatures from two closely spaced receivers (solid and dashed lines) are identical in amplitude and with only a small phase difference. The differential TEC signatures (dotted and dash-dotted lines), on the other hand, have significantly lower amplitudes, which are different between the two baselines. They are also almost in perfect antiphase. This occurs due to the difference in sign of the baseline relative to the reference at $x = 0$ km, which changes whether the measured signal leads or lags the reference signal.

The wave signature in the dTEC for a plane wave can be expressed as

$$\text{dTEC}\left(\overrightarrow{r},t\right) = \Delta\text{TEC}\left(\cos\left(\overrightarrow{k}\cdot\overrightarrow{r} - \omega t\right) - \cos(-\omega t)\right) = A\left(\Delta\text{TEC},\overrightarrow{k},\overrightarrow{r}\right)\cos\left(\phi\left(\overrightarrow{k},\overrightarrow{r}\right) - \omega t\right), \quad (1)$$

where $\overrightarrow{r}$ is the baseline from the reference station to the measurement station, t is time, ΔTEC is the wave amplitude, $\overrightarrow{k}$ is the wavevector and ω is the wave angular frequency. The amplitude $A\left(\Delta\text{TEC},\overrightarrow{k},\overrightarrow{r}\right)$ and phase $\phi\left(\overrightarrow{k},\overrightarrow{r}\right)$ of the observed wave are given by
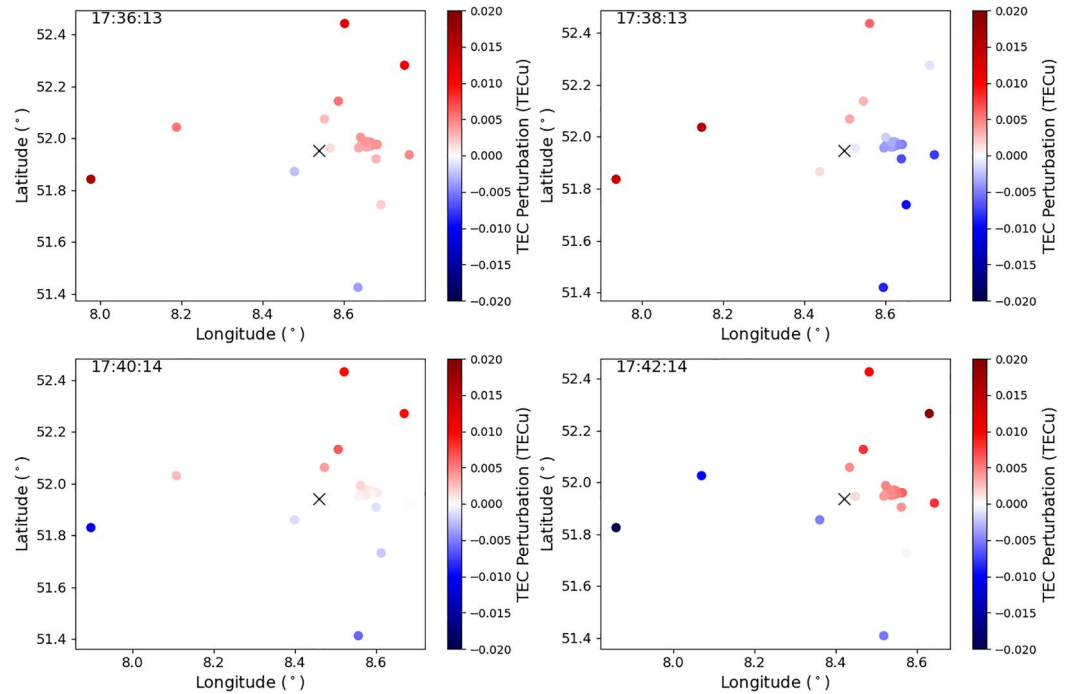
Figure 2. The detrended dTEC values for all stations, referenced to RS306 (black cross). Values are shown for four different time stamps. Points are mapped to the ionospheric pierce point location assuming a 350 km shell height. The aspect ratio is not correct causing the longitudinal distances to be exaggerated relative to the latitudinal distances.

$$A\left(\Delta\text{TEC}, \overrightarrow{k}, \overrightarrow{r}\right) = \Delta\text{TEC}\sqrt{2 - 2\cos\left(\overrightarrow{k}\cdot\overrightarrow{r}\right)}, \tag{2}$$

and

$$\phi\left(\overrightarrow{k}, \overrightarrow{r}\right) = \frac{1}{2}\left(\overrightarrow{k}\cdot\overrightarrow{r} \bmod 2\pi\right). \tag{3}$$

Naturally, the observed wave amplitude in GNSS data is independent of the measurement location for a plane wave, and the wave phase varies linearly with position. The differences between the observed wave properties for a plane wave in GNSS derived TEC and LOFAR dTEC are shown in Figure 4, based on Equations 2 and 3.

The key characteristics of the relationships for LOFAR dTEC are the low amplitude on short baselines, and the π phase discontinuity normal to $\overrightarrow{k}$ passing through the reference station. The low amplitudes on short baselines, combined with the shallower phase gradient compared to GNSS absolute TEC, lead to a practical upper limit on the wavelengths which can be detected using LOFAR dTEC. The wavelength range that can be accessed is investigated in Section 4.

3. Characterizing Waves Using Wavelet Analysis

The data shown in Section 2 illustrate that waves can be observed in the dTEC solutions, but the underlying wave parameters are not immediately apparent from the raw dTEC values. It is clear that the wavelength of the wave is larger than the longest baselines in the array as we do not see multiple wavefronts
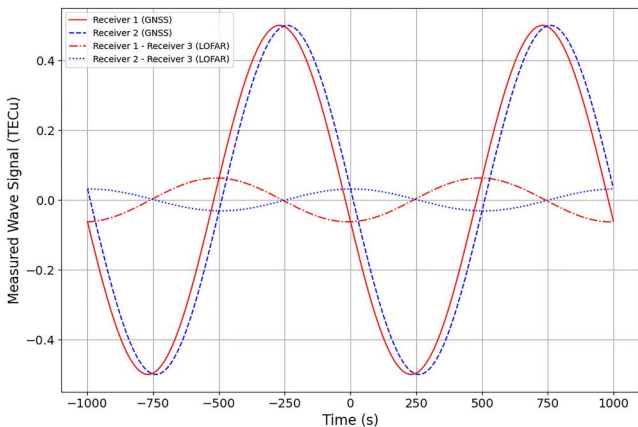

Figure 3. The expected detrended total electron content (TEC) measured from two Global Navigation Systems Satellite receivers with a single plane wave passing over them, at $x = -2$ km (Receiver 1, solid red line) and $x = 1$ km (Receiver 2, blue dashed line) respectively. Also shown are the corresponding differential TEC measurements assuming a reference receiver at $x = 0$ km (Receiver 3, red dash-dot line and blue dotted line (detrended TEC from this receiver is not shown but would be identical to those from Receivers 1 and 2 except for a small phase shift)). The wave is assumed to have a wavelength of 100 km, velocity of 100 m s^{-1} and amplitude of 0.5 TECu.
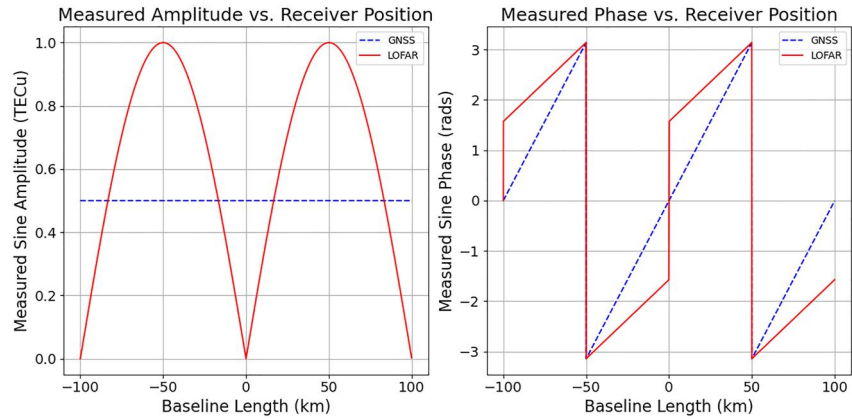
Figure 4. The measured amplitude (left) and phase (right) of a sine perturbation as a function of baseline for Global Navigation Systems Satellite receivers (blue dashed line) and LOw Frequency ARray calibration solutions (red solid line). The sine wave considered has wavelength 100 km and amplitude 0.5 TECu.

simultaneously, but the actual wavelength and amplitude of the wave are not clear. The plane wave model of Equations 2 and 3 provides a way to estimate these parameters, as it predicts a dependence of the observed amplitude and phase on baseline length and wave parameters.

The observed amplitude and phase can be estimated using a wavelet transform. This takes the observed time series and convolves it with a "wavelet" of varying time scale and time of peak amplitude. The wavelet transform therefore converts the 1D time series at each station into a complex 2D wavelet spectrum as a function of period T and time t_0, $W(T, t_0)$. The wavelet function used in this case is the Morlet wavelet, given by

$$\psi(t,s,t_0) = \sqrt{\frac{1}{s\sqrt{\pi}}} \exp\left(\frac{-(t-t_0)^2}{2s^2}\right) \exp\left(\frac{-5i(t-t_0)}{s}\right), \tag{4}$$

where t is time, s is the width parameter $\left(s(T) = \frac{5T}{2\pi}\right)$ and t_0 is the peak time. This represents a harmonic oscillation suppressed by a Gaussian, corresponding to a time limited wave signal. The wavelet coefficients are a numerical approximation to

$$W(T,t_0) = \int dTEC(t)\psi^*(t,s(T),t_0)\, dt. \tag{5}$$

This definition of the Morlet wavelet is the standard definition used by Torrence and Compo (1998) and widely implemented in software packages. However, as shown by Liu et al. (2007), it is biased toward larger periods and the amplitude of the wavelet coefficient does not correspond to the amplitude of the wave. This can be corrected by dividing the wavelet coefficients by a factor $\sqrt{f_s s}$ where f_s is the sample frequency of the time series (Liu et al., 2007).

An example wavelet power spectrum for station RS310 using RS306 as a reference is shown in Figure 5. The noise estimation was carried out using the hour of data surrounding the wave activity identified in Figure 2. The wave activity identified visually in Figure 2 corresponds to the enhancement of wavelet power at around 17:30 UT at a period of $\sim 8-10$ min.

To test the appropriateness of the simple 1-dimensional plane wave model, we first check the geographical distribution of power ($|W|^2$) and phase ($\arg(W)$) at the time and approximate period of the observed waves. This is shown in Figure 6, where the period was selected as the one with the highest power $|W|^2$ at any of the Dutch stations at this t_0. The E-W variability in both parameters is apparent, as is the discontinuity of phase and the near symmetric variation of power. This indicates that the simple plane wave model in this case provides a good approximation to the observed wave behavior and may be useful to estimate the wave parameters.
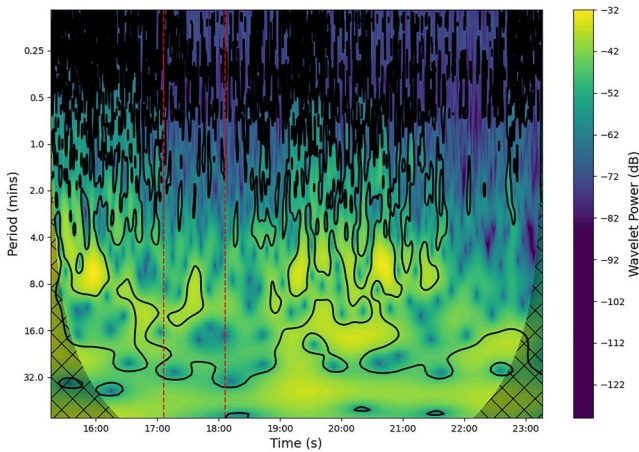
Figure 5. The wavelet power spectrum for RS310 with RS306 as the reference station. Black contours denote the 95% local significance level based on the estimated noise spectrum (see Section 3.1 for details). Red dashed lines indicate the time range over which the noise was estimated, meaning that significance levels outside this time range may be unreliable. Shaded regions at either side are the cone of influence, in which the values are unreliable due to edge effects.

3.1. Noise Estimation

As well as estimating the wave parameters, it is useful to quantify the uncertainty in the values obtained from the wavelet transform. Doing this requires an estimate of the uncertainties in the wavelet coefficients (specifically in the wavelet power and phase). This "noise" will include contributions from both instrumental/model errors in the measured values and real turbulent variations in the ionosphere, all of which can vary with time and/or observing geometry. To estimate the uncertainty in the wavelet parameters, we therefore need to estimate this noise power as a function of period and then consider the ratio of wavelet power to noise power.

In addition to providing a way to estimate uncertainty in wave power and wave phase, the noise spectrum gives a basis for specifying confidence levels for identifying signals above the noise. These are based on the fact that the noise power will be chi-square distributed with two degrees of freedom (corresponding to the real and imaginary components of the wavelet coefficient) around the underlying noise spectrum. Confidence levels can therefore be derived to determine which baselines to consider in any fitting, and exclude any which are likely to be noise dominated. The derivation of these confidence levels from the noise spectrum is explained in detail in Auchère et al. (2016).

The noise spectrum estimates rely on the Fourier power spectrum of the data to constrain them. This will in general consist of the noise spectrum with one or more peaks corresponding to coherent signals. The normalization of the Fourier power matches that of the wavelet power (before the rescaling given by Liu et al. (2007)) for noise-like signals if the Fourier power is multiplied by the length of the time series (Torrence & Compo, 1998). As a result, the noise spectrum is quantified using this normalization of the Fourier power spectrum, and then is rescaled along with the wavelet spectrum itself.

A common noise model across many areas of geophysics is the AR1 model (e.g., Torrence & Compo, 1998), which is also implemented into many common wavelet packages. This assumes that the noise can be described by a Markov process, where each value is only influenced by the preceding value. In other words, the noise contribution at timestamp i, x_i, can be expressed as

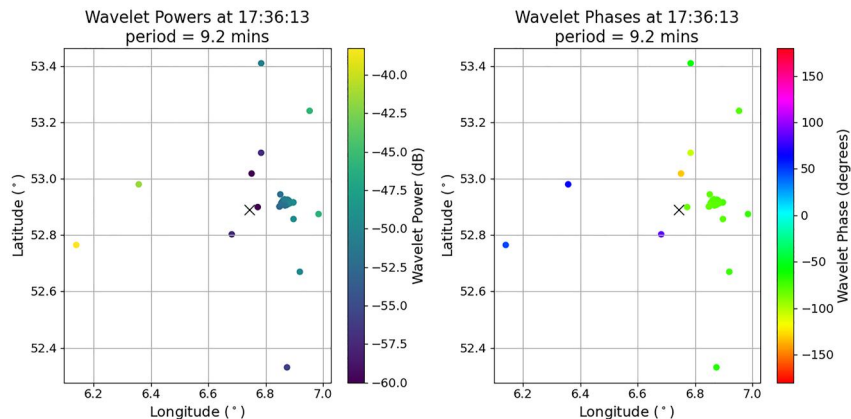$$x_i = \alpha \cdot x_{i-1} + z_i, \tag{6}$$

Figure 6. The wavelet power (left) and phase (right) calculated for a period of roughly 9.2 min at 17:36:13 UT. The black x shows the location of the reference station (RS306). The sign of the phase discontinuity is arbitrary due to the phase wrapping and an unknown offset. Aspect ratio is not set to match physical distances.

where α is the lag-one autocorrelation describing how "red" (i.e., how biased toward lower frequencies) the noise is, and z_i is a sample from a Gaussian with mean zero and variance proportional to noise power. The advantage of this is that its parameters can be very easily estimated directly from the time series, but it can only describe a limited range of spectral shapes. The dTEC data do not match the spectral shape of the AR1 model, generally having a significantly steeper spectrum, and so attempting to apply it would give extremely misleading results.

A more general approach to noise estimation is presented by Auchère et al. (2016). This involves taking the Fourier power spectrum of the data and fitting a parameterized spectrum to it. The fit is based on maximizing the likelihood using the fact that the Fourier power spectrum of a pure noise time series (assuming the noise process does not vary with time) will be distributed around the true underlying spectrum according to a chi-square with two degrees of freedom. While this is an extremely statistically thorough approach, it still requires an assumed spectral shape for the noise. We were not able to identify a suitable spectral shape for the dTEC data, as different baselines and different observations showed qualitatively different spectra.

The difficulty of identifying a suitable assumed noise spectral shape can be avoided by simply estimating the noise in a way that does not have an assumed functional form. For example, Robust Local Regression (RLR: Ruckstuhl et al., 2001, described below) makes minimal assumptions about the spectral shape. All it requires is that the noise spectrum is locally linear. While this is not the case for the dTEC data, the spectrum does locally follow a power law, meaning that it is linear in log-log space and if the fitting is carried out in this representation then the assumption of a linear spectrum holds. Another advantage of RLR is that it directly accounts for the presence of signal by iteratively excluding spikes in the spectrum from the fitting, which is not achieved by the other methods described previously.

To estimate the noise spectrum, RLR carries out weighted linear regression separately for each Fourier frequency. The weights are determined initially by the distance from each point to the point of interest in log-frequency space, with a functional form of

$$K_i = \max\!\big(0, 1 - |u_i|^3\big), \tag{7}$$

where u_i is the ratio of the separation in log-frequency space to the kernel width h. This kernel width is set for each point of interest to ensure 30 points with non-zero weight following the suggestions of Ruckstuhl et al. (2001). This provides the local component of RLR, but to ensure robustness this process must be iterated. On subsequent iterations, the local weights K_i are modified by multiplying by a robustness weight w_i given by

$$w_i = \begin{cases} \max\!\left(0, 1 - \left(\dfrac{\epsilon_i}{b\sigma}\right)^2\right), & \text{if } \epsilon_i > 0 \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

where ϵ_i is the calculated Fourier (log) power minus the fitted value from the previous iteration (i.e., the residual), b is the "robustness parameter" set to three following Ruckstuhl et al. (2001), and σ is proportional to the median absolute residual. This then works to exclude sharp (i.e. narrower than h) peaks from the fitting procedure and thereby estimate the underlying noise without being biased upwards by the signal. The estimated spectrum is found to converge sufficiently after five iterations, at which point the same weighting is used to interpolate onto the wavelet periods rather than the Fourier frequencies to give the estimated noise in the wavelet spectrum.

While fitting in log-log space is necessary to ensure a locally linear noise spectrum, it also introduces a bias into the results. This is because the mean of the logarithm of a chi-square distributed variable is lower than logarithm of its mean, but conveniently the bias can be corrected by simply adding a constant of 0.57721466 to the fitted values before rescaling from (natural) log power (Vaughan, 2005). In practice, it is also necessary to first "prewhiten" the time series before calculating the Fourier power spectrum. Otherwise, the higher frequencies will be dominated by the apparent discontinuity caused by assuming the series periodically repeats, rather than by the true noise. After fitting, the spectrum can then be corrected for this by dividing by $2 - 2\cos\!\left(\frac{2\pi f}{f_s}\right)$ (Percival & Walden, 1993). As well as the prewhitening, the signal is windowed using a Hamming window (Harris, 1978) in order to remove the effect of signal power spreading over multiple bins when the signal frequency does not match
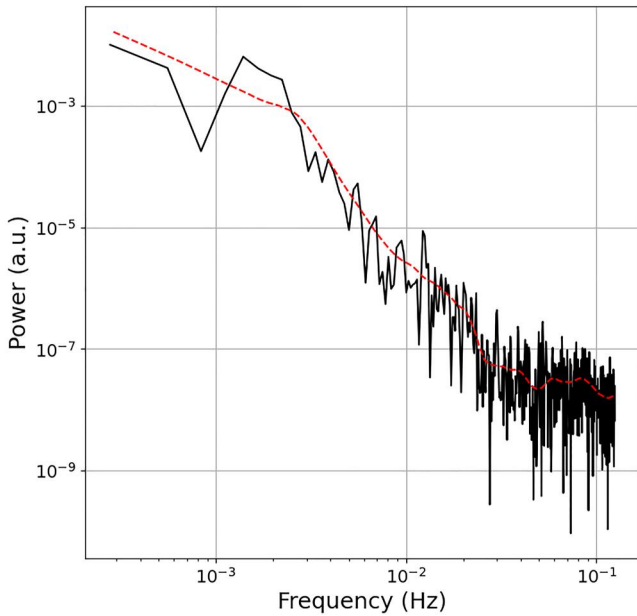
Figure 7. An example of the noise fitting procedure using an hour of the data shown in Figure 1, from 17:06 to 18:06 UT. The station used is RS310 with RS306 as the reference. The black solid line represents the Fourier power spectrum while the red dashed line is the noise spectrum estimated by RLR.

one of the Fourier frequencies. This also suppresses the noise power, creating a downward bias. To quantify this bias, the performance on artificial noise time series was tested, the results of which are described in Section 3.2.

An example of the noise fit results is shown in Figure 7, for an hour of the data shown in Figure 1 centered on the observed waves, specifically from 17:06 to 18:06 UT. The RLR method clearly does a good job of generating a smoothed version of the Fourier power spectrum as desired. There is also a distinct peak in the Fourier spectrum above the noise fit (red dashed line) between $1 \times 10^{-3} - 2 \times 10^{-3}$ Hz, corresponding to periods of around $8 - 16$ min. This is the signature of the waves visually identified in the data in Figure 2, and the RLR fit has successfully fitted under this signal as intended.

3.2. Noise Fit Validation

While the noise estimate in Figure 7 appears reasonable, given that the true noise characteristics are unknown it is useful to test the method for bias using synthetic noise. For this purpose, 10,000 1-hr long time series were generated with noise following the AR1 model of Equation 6. These included a sinusoidal signal at variable frequency and signal-to-noise ratio (SNR) to test the ability to accurately exclude any signal. The parameter α was varied between 0 and 0.99 and showed no impact on the performance of the noise estimates, as expected given no assumption is made about the exact spectral shape. These time series then had a sine wave signal added to them with variable amplitude and period, and the estimated noise was compared to the true value P_n given by (Torrence & Compo, 1998)

$$P_n = \frac{\sigma^2}{1 - 2\alpha \cos\left(\frac{2\pi t_s}{T}\right) + \alpha^2}, \tag{9}$$

where σ is the standard deviation of the sample z_i in Equation 6, t_s is the time between samples (~ 4 s) and T is the period. This comparison was made at both the signal period, and 10 times this value to capture the performance in the absence of signal.

The performance of the noise estimate for a range of signal frequencies and SNRs is shown in Figure 8. The performance is mostly uniform as shown in the left panel, with only the lowest signal frequencies showing any systematic variation. This is an unavoidable problem as for very low frequencies ($<\sim 0.8$ mHz), there are insufficient Fourier bins at lower frequencies to quantify noise power, and so the increased power due to the presence of the signal cannot be easily distinguished from an increase in noise power. However, for the vast majority of signals, this indicates a consistent performance as desired. To quantify the exact performance, the data shown in the left panel of Figure 8 were fitted by

$$\frac{P_{est}}{P_{true}} = C\left(B \exp\left(\frac{-T_0}{T}\right) + 1\right), \tag{10}$$

where P_{est} is the estimated noise power, P_{true} is the true noise power given by Equation 9 and C, B and T_0 are the fit parameters. These were found to be $C = 0.471 \pm 0.007$, $B = 570 \pm 50$ and $T_0 = 7,700 \pm 150$ s. The SNR was not included in the fit as in practice this is unknown without an accurate noise estimate, and so cannot be used to obtain the accurate noise estimate in the first place.

The rescaled noise estimates are then shown in the right panel of Figure 8. Except some at the lowest frequencies, there is no remaining trend, showing that the RLR noise estimation method with the rescaling described by Equation 10 is a generally consistent and unbiased means of estimating the noise spectrum. Applying this same correction to the estimated noise at 10 times the signal frequency leaves some residual underestimation of the noise in these regions by about 20%–25%. This is expected as a result of the one-sided robustness weighting function given by Equation 8, which inevitably underestimates noise in the absence of a signal as explained by
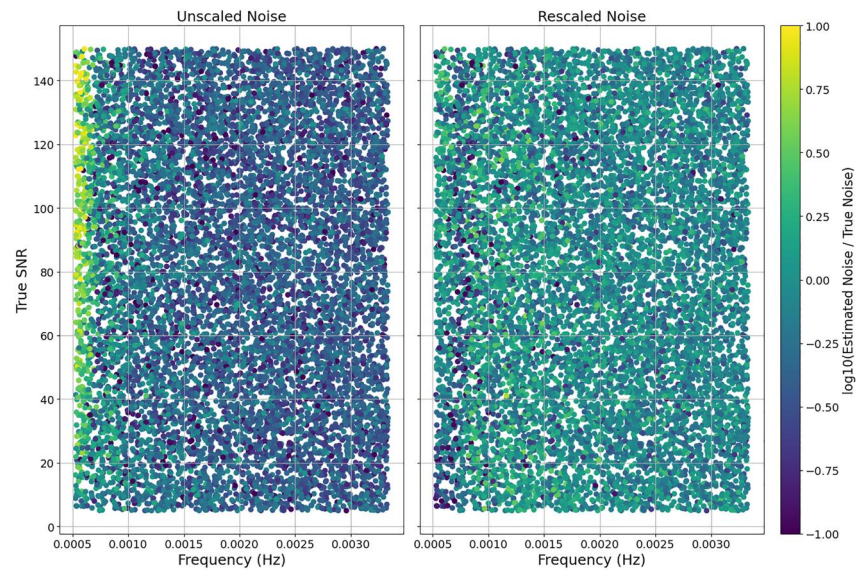
Figure 8. The errors in estimated noise power as a function of signal frequency and SNR before rescaling (left) and after rescaling by Equation 10 (right). A logarithmic color scale is used so that both overestimation and underestimation of the true noise are equally apparent.

Ruckstuhl et al. (2001). While this is not ideal, we prioritize accurate estimation of the noise in the presence of a signal, which is relevant to estimating uncertainties in the wavelet power and phase at those times, over accuracy in the absence of noise as there is no clear way to achieve both simultaneously.

3.3. Noise Correlation

Now that we can estimate the noise power, there is a second consideration in these data: the fact that the "noise" includes real turbulent variations in the ionosphere. This means that we should expect the noise on nearby baselines such as those in the LOFAR core to have appreciable correlation. Accounting for this correlation is vital in order to accurately identify both the optimal fit parameters and their uncertainty. However, measuring the spatial correlation is difficult as it requires the noise component of the signal to be isolated from genuine coherent signals, otherwise the correlation will likely be overestimated.

The correlation might also be expected to vary depending on the noise timescale considered. This is because larger timescales will generally correspond to structures with larger scale sizes which would be expected to correlate over a greater distance. Hence, it is sensible to estimate the correlation structure independently for each timescale to capture any variation resulting from this. However, larger timescales introduce the problem of the expanding "cone of influence"–the region of time to which the wavelet coefficient at a given point in time is sensitive. This reduces the number of independent samples that can be used to generate the correlation, meaning that in practice only relatively short timescales (up to ∼5 min periods for a 1 hr data set) can be analyzed in this way.

To quantify the correlation, for each baseline the noise spectrum is estimated by RLR with the rescaling described in Section 3.2. Then the wavelet spectrum is calculated, and masked to exclude all regions where the wavelet power exceeds the noise power, in order to remove, as far as possible, any contributions from coherent signals. This will also exclude some regions which are noise dominated, but robustly excluding coherent signals from consideration is the most important factor here. For each period, the correlation between both the real and the imaginary component of the wavelet coefficients on a given pair of baselines is calculated.

In this work we consider the baselines to simply be the ground level station to station baselines. The effective scan velocity of the lines of sight through the ionosphere is not considered. This will affect the apparent period of any disturbances, depending on their propagation direction and velocity. However, the apparent length scale (e.g., correlation length) and propagation direction are unaffected, and the correct period and hence velocity can be obtained if desired by calculating the line of sight scan velocity for an assumed ionospheric altitude.
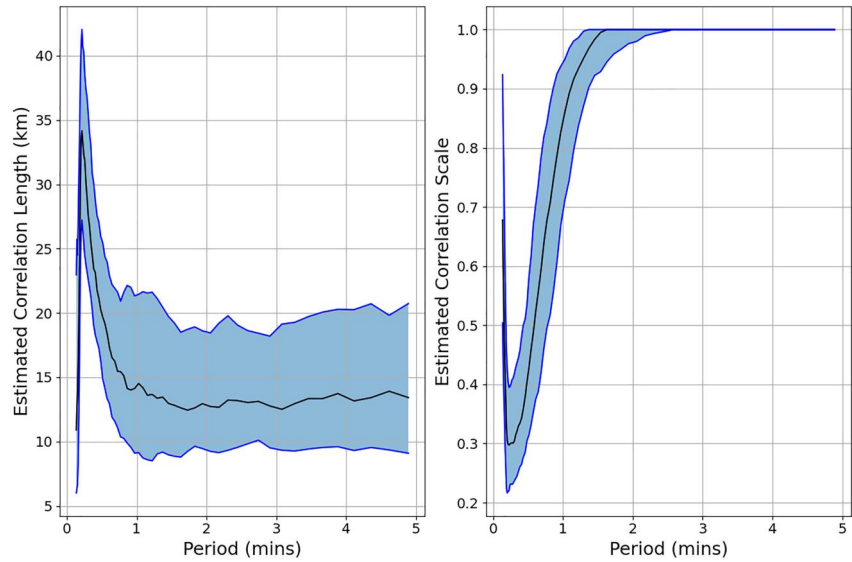
Figure 9. The median estimated correlation length (r_c, left) and correlation scale (A, right) for the simple model given by Equation 11. The shaded region indicates the interquartile range.

To represent the correlations practically, we require a simplified model of the spatial correlation function. Two options are considered: a simple exponential decay as a function of difference in baseline, and an exponential decay model assuming the noise is a combination of turbulence at both ends of the baseline. In other words, for two baselines $\vec{r_1}$ and $\vec{r_2}$, the correlation R is assumed to be described by

$$R = A \exp\left(\frac{-|\vec{r_1} - \vec{r_2}|}{r_c}\right), \tag{11}$$

where r_c is the correlation length and A is a scale factor representing the contribution of correlated errors relative to measurement errors which are assumed to be uncorrelated across baselines. This relationship is then fitted to the correlations for 100 1-hr observations taken as part of the LOFAR LBA survey (de Gasperin et al., 2021) (the observation IDs are given in the as Text S1 in Supporting Information S1). This allows us to test how stable the noise correlations are across different ionospheric conditions.

The resulting correlation length and scale factor for the simple correlation model in Equation 11 are shown in Figure 9. They show a consistent correlation length for periods over roughly 1 min. Below this, the instrumental (i.e., uncorrelated) noise starts to dominate, as shown by the lower A values, and the correlation length values themselves are largely meaningless as the correlations themselves are so low that they are essentially spurious. Averaging over the results for periods from 1 to 5 min gives a median value of 13.2 km and upper and lower quartiles of 9.2 and 19.7 km respectively. Given that this value is apparently constant between 1 and 5 min, it is reasonable to extrapolate this value to higher periods as well in the absence of a direct means of measuring correlation for the longer periods.

Although the fits here suggest $A = 1$ as the optimal solution, it was found that in practice reducing this to $A = 0.95$ was necessary. This was done because using $A = 1$ was found to give unreasonable solutions, such as predicted power exceeding all measured powers by more than 1σ. The issues with using $A = 1$ are likely due to the waves not being described precisely by an ideal plane wave, which introduces another source of variation not captured by the estimated covariance. The mean square residuals on the fit also increase with increasing period, although this may reflect the reducing effective sample size at longer periods and hence higher uncertainty in the measured correlations themselves than any reduction in the accuracy of the model itself in describing the relationship. For all further analysis, r_c was taken to be the median value of 13.2 km.
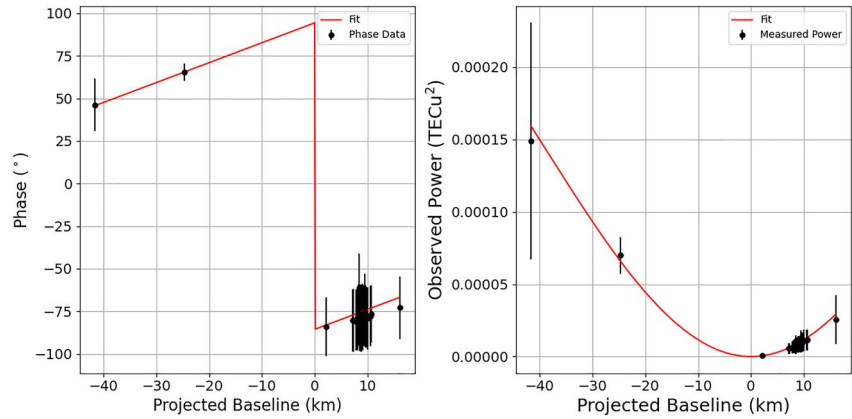
Figure 10. The phase (left) and power (right) data as a function of projected baseline length with the fitted values shown. These correspond to the same T and t_0 as the values shown in Figure 6. The fit parameters were: $|\vec{k}| = 4.08 \pm 0.94 \times 10\text{-}5\,\text{m}-1$, $\theta = 84.4 \pm 2.8°$, and $\Delta\text{TEC} = 8.4 \pm 1.9$ mTECu.

3.4. Fit Method

Using the wavelet coefficients, estimated noise power and estimate of the noise correlation, we can therefore define a covariance matrix corresponding to our data and use this to define the best fit parameters. Only baselines for which the power exceeds the 95% local confidence level (approximately 3 times the noise power) are included in the fitting. The best fit parameters are those which minimize

$$\chi^2 = \vec{\epsilon} \cdot \left(\underline{\underline{C}}^{-1} \cdot \vec{\epsilon} \right), \tag{12}$$

where $\vec{\epsilon}$ is the vector of phase and power residuals after fitting to Equations 2 and 3 and $\underline{\underline{C}}^{-1}$ is the inverse of their combined covariance matrix. The residual vector $\vec{\epsilon}$ is determined by the four free parameters of the fit: the wavevector magnitude $|\vec{k}|$, the wavevector azimuth θ, the phase offset ωt and the wave amplitude ΔTEC. The fit is carried out for both phase and power simultaneously as both relationships are dependent on wavevector, and so fitting them in parallel is required to obtain a fully accurate optimum solution.

To ensure the fitting algorithm reliably finds the optimum solution, it must be provided with a reasonable initial estimate of the parameters. All components of this rely on the assumption that the wavelength of the ionospheric wave is significantly greater than the baseline lengths, so that there is only one phase discontinuity and the power variation is approximately quadratic. The first parameter to be estimated is the azimuth, using the wavelet powers and following the approach of Beser et al. (2022). We define a quantity $\frac{|W|^2}{|r|^2}\vec{r}$ which should in principle be independent of baseline length and only depend on baseline orientation relative to the wave propagation direction. We then use principal component analysis (PCA) to identify the dominant direction in this quantity, which ideally corresponds to the wave direction. Due to the uneven distribution of baseline directions provided by LOFAR, this initial estimate has a bias which needs to be corrected. This is achieved by generating artificial data at a range of azimuths and using the PCA estimated azimuths as a basis for inverting back to the true azimuth, described in Appendix A.

With an estimated azimuth, we can then project the baselines onto this direction to estimate the other parameters, giving projected baselines x. An example of this projection is shown in Figure 10, although in this case it is projected onto the final optimal fitted azimuth rather than the initial PCA estimate. We sort the phases by projected baseline and unwrap them, then find the sign of the discontinuity and remove it from the sorted data. In Figure 10, this would mean adding 180° to all phases with $x > 0$. After removing the discontinuity, the phase should be linear with projected baseline, and so we use linear regression to provide an estimate for $|\vec{k}|$ and ωt. With the estimate of $|\vec{k}|$ we then estimate ΔTEC by averaging $\frac{|W|^2}{\left(|\vec{k}|_x \right)^2}$ on the longest positive and negative projected baselines, which is the long wavelength approximation to Equation 2.

While in most cases the method as described above provides a sufficiently good initial estimate to find the optimum solution, it can encounter problems in some cases due to the discontinuity in the phase relationship. This can cause very sharp changes in χ^2 with θ where the sign of one projected baseline x changes and hence moves to the other side of the phase discontinuity, which can be local minima and hence confuse the fit. To mitigate this issue, after the fit, any baselines which are within $\max(0.1°, \sigma_\theta)$ (σ_θ is the estimated uncertainty in azimuth) of being perpendicular to the wavevector are removed and the fit is repeated, using the previous "optimal" solution as its initial guess. The floor of 0.1° is necessary as in some cases the sharp change in χ^2 leads to an estimated σ_θ which is smaller than the step size used in the fit process. Once no baselines lie within $\max(0.1°, \sigma_\theta)$ of perpendicularity to $\vec{k}$, the fit is accepted.

Figure 10 shows the final fits to the phase and power for the sample data, using the same T and t_0 as were shown in Figure 6. The agreement between the simple plane wave model and the data is very clear, but the uncertainties seem to significantly overestimate the spread of the data in this case. The fit suggests a wavelength of 154 ± 35 km, which is toward the lower end of what is typically classified as an MSTID. The estimated amplitude of 8.4 ± 1.9 mTECu is significantly below the level which is typically required for confident identification of TIDs in GNSS TEC maps (e.g., Otsuka et al. (2013) used a threshold of 0.2 TECu) and around the sensitivity threshold of GNSS measurements, but well above the noise floor of LOFAR dTEC. This demonstrates the potential of LOFAR dTEC measurements to reveal waves which have previously not been detectable due to the lower noise floor (Mevius et al., 2016).

4. Validation Using Synthetic Data

While the performance of the method on the sample data is promising, there is no way to directly verify that the derived values are accurate for real data. To account for this issue, synthetic data can be generated, meaning that the true wave parameters are known and can be compared to the derived estimates. This allows the performance of the wavelet method to be investigated across a range of parameters, identifying any bias in azimuth and the range of wavelengths that can be accurately identified. The residuals of the wavelet phase and power to their known true values can also be compared to the estimated uncertainties to confirm that the noise estimation performs as expected. The performance on synthetic data will inevitably be better than for real data as the assumption of a plane wave is exact for the synthetic data.

The noise is replicated by a simplified model, the AR1 model discussed in the previous section with $\alpha = 0.95$. While this is not an accurate representation of the noise in the real data, given that the fit method is not dependent on the spectral shape it should be suitable for this purpose. The random samples (z_i in Equation 6) are generated with correlation between baselines to replicate the correlation in the data (assumed $r_c = 13.2$ km, $A = 0.95$). This means that the assumed correlation used to generate $\underline{C}$ in Equation 12 is exactly matched to the true correlation for the synthetic data. As a result, the performance of the fit method on the synthetic data should be better than its performance on real data where the correlations are only approximate.

One further possible characteristic of the noise in the real data that is not replicated here is that it may be directly driven by the waves themselves. The density gradients associated with the wave can favor the development of the gradient-drift instability and resulting turbulent cascade (e.g., Lin et al., 2016). This would mean that the noise would be enhanced in the presence of a wave relative to the background. However, this is not investigated further here.

To improve the fit performance, it proved to be necessary to attempt the fit using several different references to get good performance across all azimuths. A set of three references were found to be sufficient for this, namely RS205, RS306, RS406. These provide a range of orientations relative to the core, while also ensuring that there are a reasonable number of baselines across a wide range of directions (i.e., these references are not on the edge of the array). This is important because if the selected reference is on the edge of the array, the phase discontinuity will not be present which significantly reduces the ability to identify the propagation direction. The selected fit parameters were then those from the reference which gave the smallest value of χ^2 in Equation 12.

The synthetic data generation creates 1 hr of data per set of wave parameters. This is based on the typical 1 hr duration of individual observing windows in LOFAR surveys (e.g., de Gasperin et al., 2021). The wave parameters are generated using Sobol sequences (Saltelli et al., 2010; Sobol & Levitan, 1999) to generate representative sampling across the parameter space with minimum clustering. The parameters generated are as
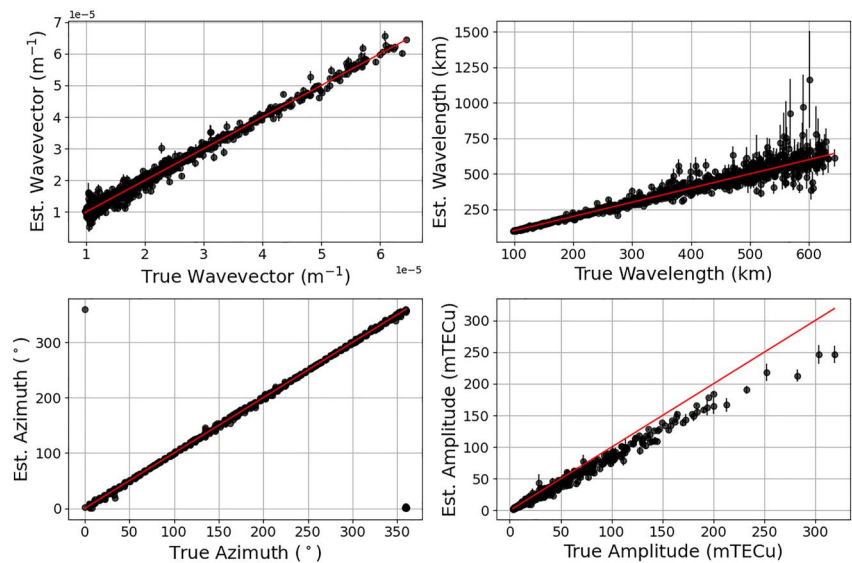
Figure 11. The relationship between true (horizontal axis) and estimated (vertical axis) parameters. Top left is wavevector $|\vec{k}|$, top right is wavelength Λ, bottom left is azimuth θ and bottom right is amplitude ΔTEC. Red lines represent the ideal (i.e., $x = y$) behavior.

follows: wavelength Λ, between 100 and 625 km; azimuth θ, between 0 and 360°; target median SNR, between 5 and 50; and a "period scale," between 0 and 1. The period scale is used to ensure that the wave velocity is physically reasonable for MSTIDs, with a maximum velocity set at 400 m s^{-1}. Based on the generated wavelength, the minimum period is then defined as the period which would give a velocity of 400 m s^{-1}, and then the period scale determines the wave period linearly between this minimum period and the maximum period set by the wavelet cone of influence, which is ~26 min. The result of this generation scheme is that wavelength is sampled effectively uniformly, while periods are biased toward the higher end of the available range and amplitudes are roughly proportional to wavelength with the proportionality constant determined by the target SNR. The minimum wavelength of 100 km is set to ensure no additional phase discontinuities occur within the size of the array. While the fit method can in theory account for these, it will likely impact performance and so is not considered for this analysis.

The correlation between true and estimated values of the TID parameters are shown in Figure 11. These plots exclude any fit for which the estimated uncertainty in $|\vec{k}|$ or ΔTEC exceeds the fit value (i.e., relative uncertainty >100%), and those for which the estimated wavelength was over 5,000 km, which collectively accounted for only 0.4% of fits. Statistical measures of the fit performance are also given in Table 1, indicating the high correlation of all parameters. The root mean square error (RMSE) in wavelength Λ is high, but this is largely dominated by the decrease in performance at long wavelengths shown in Figure 11. At shorter wavelengths, the errors will be much smaller than ~50 km. The other clear issue is a systematic underestimation of amplitude ΔTEC. This is a major contributor to the RMSE on ΔTEC, and, if the data are rescaled to remove the bias, the RMSE reduces to 4.25 mTECu. The azimuth is roughly uniformly accurate at all values, with only a slight decrease in accuracy for roughly N-S propagation ($\theta \sim 0°$ or 180°) compared to E-W propagation visible in Figure 11. The reason for this is likely the shorter E-W baselines available compared to N-S baselines, making the data less sensitive to changes in azimuth for N-S propagation.

The bias in amplitude is partially explained by a previously unreported bias in the wavelet transform itself. Applying the wavelet transform to a sine wave of unit amplitude, the amplitude of the wavelet coefficient at the corresponding frequency is roughly 0.93. This was not reported by Liu et al. (2007), likely as that work was focused on removing the bias in comparing amplitudes across

Table 1
The Performance of the Fit in Predicting the Various Wave Parameters

Parameter	RMSE	Bias	Correlation
$|\vec{k}|$	1.33×10^{-6} m^{-1}	-0.03%	0.994
Λ	45.8 km	0.03%	0.959
ΔTEC	8.56 mTECu	-12.86%	0.994
θ	1.697°	$-0.065°$	1.000

Note. RMSE is root mean square error. Bias is median percentage error except for θ where it is median error.
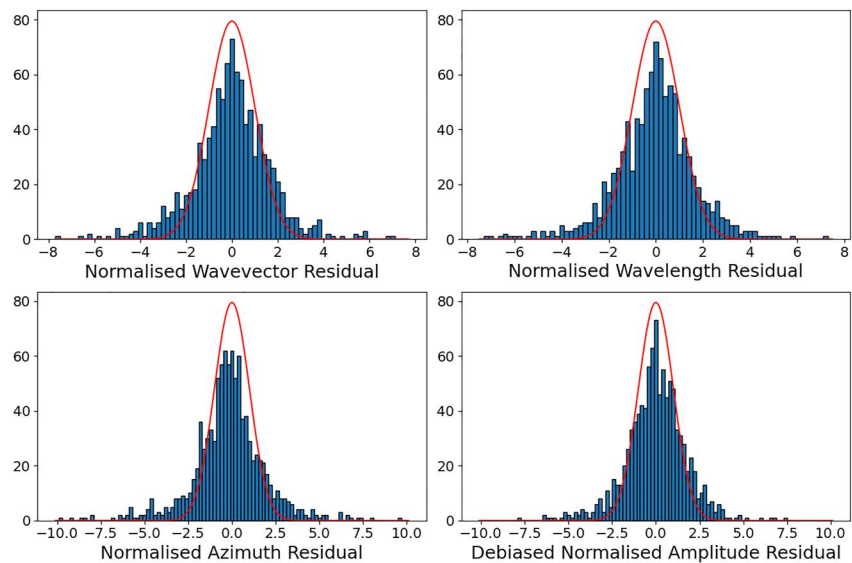
Figure 12. The normalized residuals of the fitted parameters relative to the true values. Amplitude estimates are first debiased by rescaling according to the median percentage error in each 200 s period bin. The vertical scale gives the number of samples in each bin. Normalization consists of dividing the residual by the estimated uncertainty. Red curves show a unit Gaussian for reference.

scales rather than accurately determining the amplitude itself. This would give rise to a $\sim -7\%$ bias however, not the -12.86% found here. An extra contribution may be that at longer periods, even points in the wavelet spectrum which are outside the cone of influence (the region in which the effects of the finite length of the time series are generally defined as significant) still show noticeable decreases in amplitude beyond the $\sim -7\%$ general bias. The bias toward the longer periods in the sampling method used to generate the synthetic data is therefore likely responsible for this further underestimation of the true amplitude.

The other notable feature in Figure 11 is the increased spread around the true value at longer wavelengths/ smaller wavevectors. This represents a fundamental limitation of LOFAR in identifying large scale waves. The wavelength is primarily determined from the gradient of the phase in Equation 3, which decreases as wavelength increases. When the wavelength is significantly larger than the largest baselines in the array (~ 100 km), the uncertainty in the phase values will significantly degrade the estimated gradient and hence wavelength. The results here suggest that this decrease in performance is significant for wavelengths above roughly $400-500$ km for the moderate to high SNR values considered here. At low SNR, this effect would likely be even more severe.

As well as the performance of the estimated wave parameters themselves, it is useful to consider how well the uncertainties in these parameters were estimated. Figure 12 shows the normalized residuals (i.e., (estimated value–true value)/estimated uncertainty) for the various fit parameters. In order to accurately debias the amplitude residuals, it proved necessary to calculate the median percentage error independently in 200 s period bins from 300 to 1,700 s. The resulting values were consistent with the explanation proposed above for the bias, with the lower period bins showing biases of $\sim -6.5\%$ before dropping significantly above ~ 900 s, reaching -21.8% in the highest bin. By comparison to the unit Gaussians (shown in red), all uncertainties seem to be slightly underestimated. The standard deviations of the normalized residuals in each parameter range from ~ 1.68 to 2.15, where a standard deviation of 1 would be expected for perfect uncertainty estimation.

Some of this underestimation of the errors may be related to mismatching of the true wave period to the wavelet period, as the fit relies on the closest wavelet period. Another factor is the imperfect nature of the noise estimation method. While it is unbiased (after rescaling), there is still appreciable variance of the estimated noise value around the true value. This will lead to the fit giving more weight than it should to certain baselines and less to others, making the estimated parameters less accurate in a way which cannot be captured by the uncertainty estimates based on those noise estimates. A small improvement in the uncertainty estimates can be

achieved by multiplying the uncertainty by the square root of the reduced chi-square χ_r^2 of the final fit. This is defined as

$$\chi_r^2 = \frac{\chi^2}{2N_{\text{base}} - 4},$$ (13)

where χ^2 is as defined in Equation 12, N_{base} is the number of baselines included in the fit and four is the number of fit parameters. This rescaling is equivalent to ignoring the absolute value of the different uncertainty estimates and only considering their relative magnitudes. This rescaling provides standard deviations of $\sim 1.47 - 1.72$ for the normalized residuals. If this rescaling were applied to the example shown in Figure 10, the results would be: $\Lambda = 154 \pm 8$ km, $\Delta \text{TEC} = 8.4 \pm 0.4$ mTECu, $\theta = 84.4 \pm 0.7°$.

5. Discussion and Conclusions

The dTEC values calculated for calibration of astronomical observations with LOFAR represent a largely unexplored source of data on ionospheric dynamics over Europe. They are obtained with extremely high precision (Mevius et al., 2016), and there is already a large amount of data available thanks to the regular astronomical survey observations made using LOFAR (e.g., de Gasperin et al., 2021; Heald et al., 2015; Shimwell et al., 2017). These data can enable more detailed studies of TIDs over Europe, including lower amplitude and shorter wavelength waves, and therefore giving a more complete picture of the dynamics of the ionosphere and, ultimately, the neutral atmosphere in which it is embedded.

This paper presents a method for extracting information on waves present in the data, and it is shown based on artificial data that waves with wavelengths up to ~ 500 km (roughly 5–10 times the size of the array) can be reliably identified. Larger wavelengths may also be identified but this is only possible for high amplitudes as the SNR required increases with increasing wavelength. In a future paper, this method will be used to characterize wave behavior above LOFAR on a statistical basis.

The approach developed here is also not highly specific to LOFAR, and could be adapted to identifying waves in data from other distributed networks. The specific fitted relationships would naturally differ in the case of absolute rather than differential measurements. However, the general method of quantifying the noise spectra in particular may be useful for a range of applications. This is especially true if the underestimate in parameter uncertainties shown in Figure 12 can be corrected.

As shown in Section 3.4, waves are detectable in LOFAR dTEC which would be below the detection threshold for GNSS TEC. This suggests that these data can complement existing work based on analysis of GNSS data (e.g., Otsuka et al., 2013) by extending the range of wave parameter space that can be sampled, particularly to lower amplitudes and possibly to shorter wavelengths as well. It also complements work using scintillation data from LOFAR which characterizes TID structure on scales of ~ 20 km (e.g., Boyde et al., 2022; Dorrian et al., 2023).

Given the large number of very short (100s of meters to kilometres) baselines within the LOFAR core, a similar analysis method to that presented here but applied only to the core stations may be able to probe even shorter wavelengths than those $> \sim 100$ km considered in this paper. However, this would be dependent on whether the correlation length shown in Figure 9 poses an obstacle to comparing signals between such similar baselines, which is not considered in this work. It may also be possible to identify such short wavelength waves using the full Dutch network, but the fit method will likely lose reliability when additional phase discontinuities become present, and distortions from the simple plane wave model will likely be more significant over multiple wavelengths. Similarly, including the international stations may enable longer wavelength waves to be identified, extending above the ~ 500 km upper wavelength limit identified here. However, as the longest baselines are in the E-W direction and most LSTIDs propagate equatorwards from the auroral zone (e.g., Habarulema et al., 2013), this may not provide the most favorable baseline distribution.

Appendix A: PCA Bias Correction

The PCA approach to estimating propagation direction is accurate in the idealized case only with a roughly uniform distribution of baselines in azimuth. With LOFAR this is not the case, and as a result the estimated azimuth will be biased. This bias is non-uniform in azimuth, and dependent on the reference station selected as this determines the azimuth distribution of the baselines.
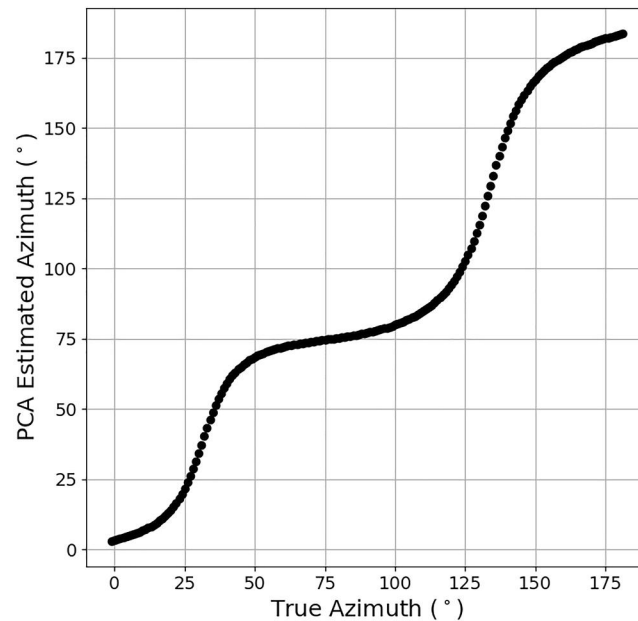

Figure A1. The bias in the azimuth estimated from PCA using RS306 as a reference in the idealized case.

In order to correct for this bias, for a given set of baselines, synthetic data corresponding to the idealized case (i.e., power proportional to the baseline projected onto the propagation direction squared) can be generated for a range of azimuths. For each input azimuth, the estimated azimuth from the PCA is then determined. An example of the relationship obtained using RS306 as a reference is shown in Figure A1. This displays clear "preferred directions" around $70 - 85°$ and $-5 - 10/175 - 190°$ where many true azimuths generate estimated azimuths in the same narrow ranges.

Using this idealized picture of the bias, it is generally possible to define a means of inverting the estimates to mitigate the effect of the bias. This is possible provided the variation is monotonic, as it is in Figure A1. In this case, a correction function can be defined as a linear interpolation of the relationship between true azimuth and estimated azimuth, and used to map the estimates back to something closer to their true value. Neglecting such a correction has been found to reduce reliability of the fitting, and in some cases prevent a viable solution from being found, even when the wave activity is clear from a visual inspection of the data.

Data Availability Statement

All data used in this paper were obtained using LOFAR. Data from LOFAR is available from the LOFAR Long Term Archive at https://lta.lofar.eu/.

References

Aarons, J. (1982). Global morphology of ionospheric scintillations. *Proceedings of the IEEE*, *70*(4), 360–378. https://doi.org/10.1109/PROC.1982.12314

Auchère, F., Froment, C., Bocchialini, K., Buchlin, E., & Solomon, J. (2016). On the Fourier and wavelet analysis of coronal time series. *The Astrophysical Journal*, *825*(2), 110. https://doi.org/10.3847/0004-637X/825/2/110

Beser, K., Mevius, M., Grzesiak, M., & Rothkaehl, H. (2022). Detection of periodic disturbances in LOFAR calibration solutions. *Remote Sensing*, *14*(7), 1719. https://doi.org/10.3390/rs14071719

Boyde, B., Wood, A., Dorrian, G., Fallows, R. A., Themens, D., Mielich, J., et al. (2022). Lensing from small-scale travelling ionospheric disturbances observed using LOFAR. *Journal of Space Weather and Space Climate*, *12*, 34. https://doi.org/10.1051/swsc/2022030

Carbone, D., van der Horst, A. J., Wijers, R. A. M. J., Swinbank, J. D., Rowlinson, A., Broderick, J. W., et al. (2016). New methods to constrain the radio transient rate: Results from a survey of four fields with LOFAR. *Monthly Notices of the Royal Astronomical Society*, *459*(3), 3161–3174. https://doi.org/10.1093/mnras/stw539

de Gasperin, F., Dijkema, T. J., Drabent, A., Mevius, M., Rafferty, D., van Weeren, R., et al. (2019). Systematic effects in LOFAR data: A unified calibration strategy. *A&A*, *622*, A5. https://doi.org/10.1051/0004-6361/201833867

de Gasperin, F., Mevius, M., Rafferty, D., Intema, H., & Fallows, R. (2018). The effect of the ionosphere on ultra-low-frequency radio-interferometric observations. *Astronomy & Astrophysics*, *615*, A179. https://doi.org/10.1051/0004-6361/201833012

de Gasperin, F., Williams, W. L., Best, P., Brüggen, M., Brunetti, G., Cuciti, V., et al. (2021). The LOFAR LBA sky survey-I. Survey description and preliminary data release. *Astronomy & Astrophysics*, *648*, A104. https://doi.org/10.1051/0004-6361/202140316

Dorrian, G., Fallows, R., Wood, A., Themens, D. R., Boyde, B., Krankowski, A., et al. (2023). LOFAR observations of substructure within a traveling ionospheric disturbance at mid-latitude. *Space Weather*, *21*(1), e2022SW003198. https://doi.org/10.1029/2022SW003198

Fallows, R. A., Forte, B., Astin, I., Allbrook, T., Arnold, A., Wood, A., et al. (2020). A LOFAR observation of ionospheric scintillation from two simultaneous travelling ionospheric disturbances. *Journal of Space Weather and Space Climate*, *10*, 10. https://doi.org/10.1051/swsc/2020010

Fritts, D. C., & Alexander, M. J. (2003). Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, *41*(1). https://doi.org/10.1029/2001RG000106

Habarulema, J. B., Katamzi, Z. T., & McKinnell, L.-A. (2013). Estimating the propagation characteristics of large-scale traveling ionospheric disturbances using ground-based and satellite data. *Journal of Geophysical Research: Space Physics*, *118*(12), 7768–7782. https://doi.org/10.1002/2013JA018997

Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, *66*(1), 51–83. https://doi.org/10.1109/PROC.1978.10837

Heald, G. H., Pizzo, R. F., Orrú, E., Breton, R. P., Carbone, D., Ferrari, C., et al. (2015). The LOFAR multifrequency snapshot sky survey (MSSS)-I. Survey description and first results. *Astronomy and Astrophysics*, *582*, A123. https://doi.org/10.1051/0004-6361/201425210

Helmboldt, J., Lazio, T., & Dymond, K. (2012). A new technique for spectral analysis of ionospheric TEC fluctuations observed with the Very Large Array VHF system: From QP echoes to MSTIDs. *Radio Science*, *47*(04), 1–21. https://doi.org/10.1029/2011RS004787

Hines, C. O. (1960). Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics*, *38*(11), 1441–1481. https://doi.org/10.1139/p60-150

Hocke, K., & Schlegel, K. (1996). A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982–1995. *Annales Geophysicae*, *14*(9), 917–940. https://doi.org/10.1007/s00585-996-0917-6

Hoogeveen, G., & Jacobson, A. (1997). Improved analysis of plasmasphere motion using the VLA radio interferometer. *Annales Geophysicae*, *15*(2), 236–245. https://doi.org/10.1007/s00585-997-0236-6

Hoque, M. M., & Jakowski, N. (2008). Estimate of higher order ionospheric errors in GNSS positioning. *Radio Science*, *43*(05), 1–15. https://doi.org/10.1029/2007RS003817

Hunsucker, R. D. (1982). Atmospheric gravity waves generated in the high-latitude ionosphere: A review. *Reviews of Geophysics*, *20*(2), 293–315. https://doi.org/10.1029/RG020i002p00293

Jacobson, A. R., & Erickson, W. C. (1992). Wavenumber-resolved observations of ionospheric waves using the Very Large Array radiotelescope. *Planetary and Space Science*, *40*(4), 447–455. https://doi.org/10.1016/0032-0633(92)90163-I

Jacobson, A. R., & Erickson, W. C. (1993). Observations of electron density irregularities in the plasmasphere using the VLA radio interferometer. *Annales Geophysicae*, *11*(10), 869–888.

Lan, J., Ning, B., Li, G., Zhu, Z., Hu, L., & Sun, W. (2018). Observation of short-period ionospheric disturbances using a portable digital ionosonde at Sanya. *Radio Science*, *53*(12), 1521–1532. https://doi.org/10.1029/2018RS006699

Lin, F. F., Wang, C. Y., Su, C. L., Shiokawa, K., Saito, S., & Chu, Y. H. (2016). Coordinated observations of F region 3 m field-aligned plasma irregularities associated with medium-scale traveling ionospheric disturbances. *Journal of Geophysical Research: Space Physics*, *121*(4), 3750–3766. https://doi.org/10.1002/2016JA022511

Liu, Y., San Liang, X., & Weisberg, R. H. (2007). Rectification of the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2093–2102. https://doi.org/10.1175/2007JTECHO511.1

Mevius, M., van der Tol, S., Pandey, V. N., Vedantham, H. K., Brentjens, M. A., de Bruyn, A. G., et al. (2016). Probing ionospheric structures using the LOFAR radio telescope. *Radio Science*, *51*(7), 927–941. https://doi.org/10.1002/2016RS006028

Munro, G. H. (1950). Travelling disturbances in the ionosphere. *Proceedings of the Royal Society of London - Series A: Mathematical and Physical Sciences*, *202*(1069), 208–223.

Otsuka, Y., Suzuki, K., Nakagawa, S., Nishioka, M., Shiokawa, K., & Tsugawa, A. (2013). GPS observations of medium-scale traveling ionospheric disturbances over Europe. *Annales Geophysicae*, *31*(2), 163–172. https://doi.org/10.5194/angeo-31-163-2013

Pearson, T., & Readhead, A. (1984). Image formation by self-calibration in radio astronomy. *Annual Review of Astronomy and Astrophysics*, *22*(1), 97–130. https://doi.org/10.1146/annurev.astro.22.1.97

Percival, D. B., & Walden, A. T. (1993). Parametric spectral estimation. In *Spectral analysis for physical applications* (pp. 391–455). Cambridge University Press. https://doi.org/10.1017/CBO9780511622762.012

Ruckstuhl, A. F., Jacobson, M. P., Field, R. W., & Dodd, J. A. (2001). Baseline subtraction using robust local regression estimation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *68*(2), 179–193. https://doi.org/10.1016/S0022-4073(00)00021-2

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, *181*(2), 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Shimwell, T. W., Röttgering, H. J. A., Best, P. N., Williams, W. L., Dijkema, T. J., de Gasperin, F., et al. (2017). The LOFAR two-metre sky survey-I. survey description and preliminary data release. *Astronomy & Astrophysics*, *598*, A104. https://doi.org/10.1051/0004-6361/201629313

Sobol, I., & Levitan, Y. L. (1999). A pseudo-random number generator for personal computers. *Computers & Mathematics with Applications*, *37*(4–5), 33–40. https://doi.org/10.1016/S0898-1221(99)00057-7

Themens, D. R., Watson, C., Žagar, N., Vasylkevych, S., Elvidge, S., McCaffrey, A., et al. (2022). Global propagation of ionospheric disturbances associated with the 2022 Tonga volcanic eruption. *Geophysical Research Letters*, *49*(7), e2022GL098158. https://doi.org/10.1029/2022GL098158

Thompson, A. R., Clark, B., Wade, C., & Napier, P. J. (1980). The very large array. *The Astrophysical Journal - Supplement Series*, *44*, 151–167. https://doi.org/10.1086/190688

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079⟨0061:APGTWA⟩2.0.CO;2

van de Kamp, M., Pokhotelov, D., & Kauristie, K. (2014). TID characterised using joint effort of incoherent scatter radar and GPS. *Annales Geophysicae*, *32*(12), 1511–1532. https://doi.org/10.5194/angeo-32-1511-2014

van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., et al. (2013). LOFAR: The LOw-Frequency ARray. *Astronomy and Astrophysics*, *556*, A2. https://doi.org/10.1051/0004-6361/201220873

Vaughan, S. (2005). A simple test for periodic signals in red noise. *Astronomy & Astrophysics*, *431*(1), 391–403. https://doi.org/10.1051/0004-6361:20041453

Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., Labropoulos, P., Pandey, V. N., Kazemi, S., et al. (2013). Initial deep LOFAR observations of epoch of reionization windows-I. the north celestial pole. *Astronomy & Astrophysics*, *550*, A136. https://doi.org/10.1051/0004-6361/201220874