Scale-Separated Dynamic Mode Decomposition and Ionospheric Forecasting

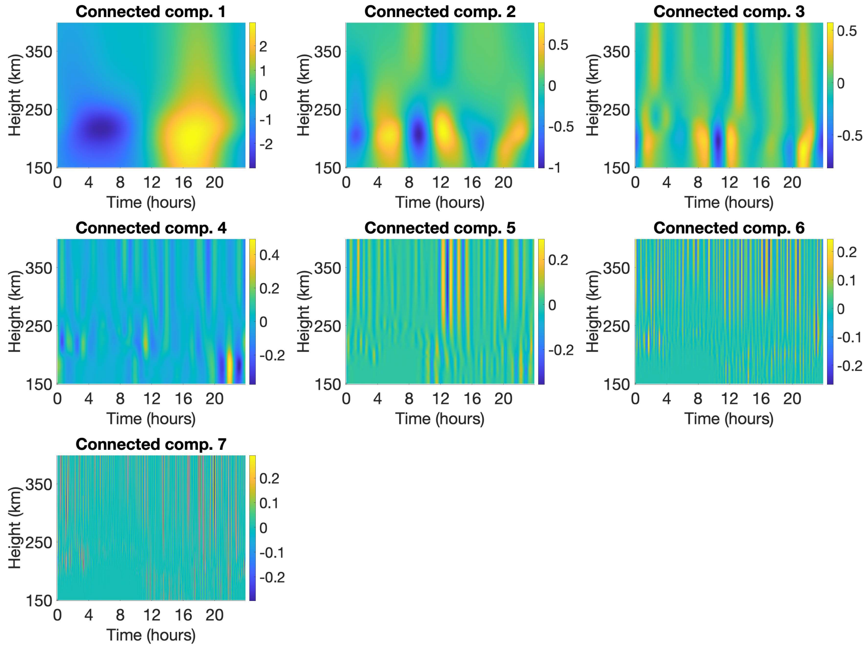
Daniel Jay Alford-Lago[1], Christopher W Curtis[2], Alexander Ihler[3], and Katherine Anne Zawdie[4]

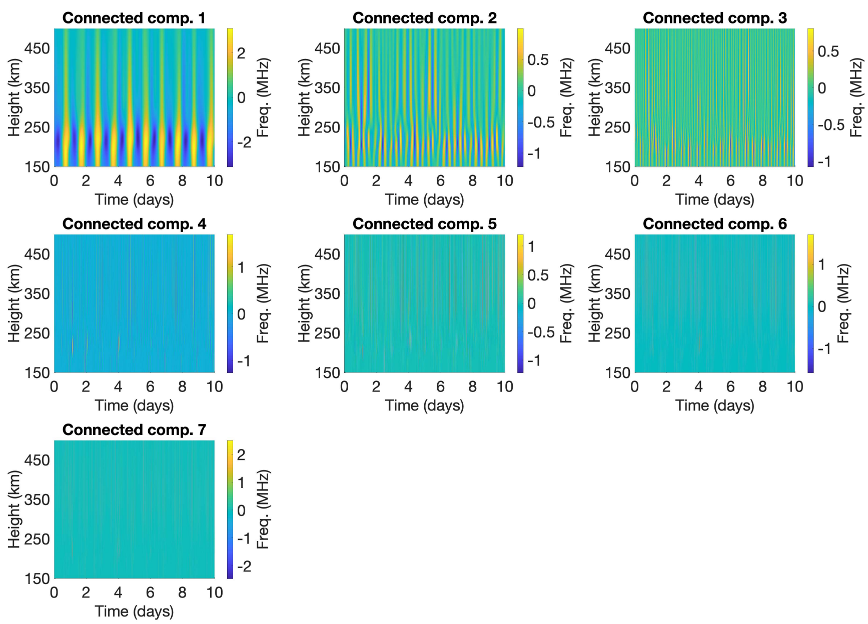
[1]Naval Information Warfare Center Pacific
[2]San Diego State University
[3]UC Irvine
[4]Naval Research Laboratory

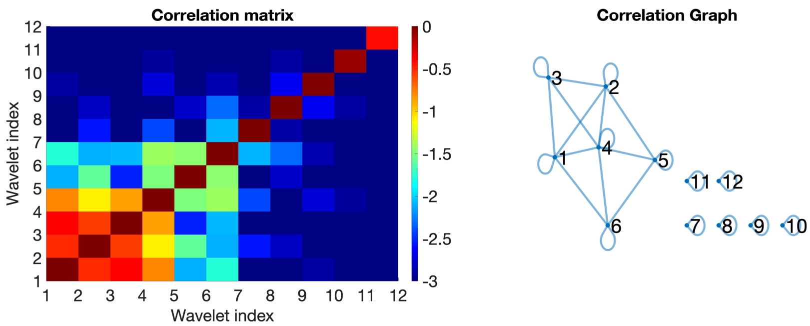
December 22, 2022

Abstract

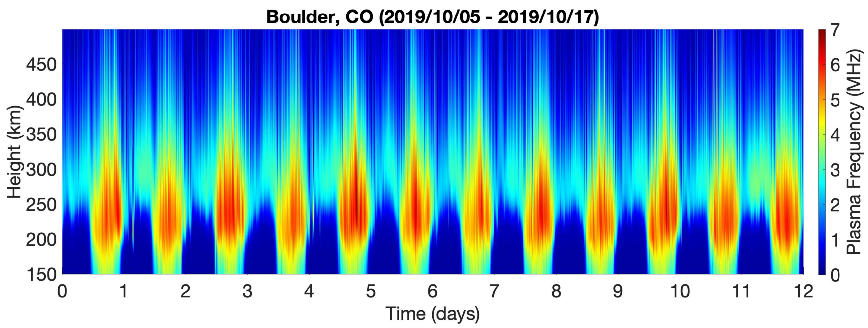
We present a method for forecasting the foF2 and hmF2 parameters using modal decompositions from measured ionospheric electron density profiles. Our method is based on Dynamic Mode Decomposition (DMD), which provides a means of determining spatiotemporal modes from measurements alone. Our proposed extensions to DMD use wavelet decompositions that provide separation of a wide range of high-intensity, transient temporal scales in the measured data. This scale separation allows for DMD models to be fit on each scale individually, and we show that together they generate a more accurate forecast of the time-evolution of the F-layer peak. We call this method the Scale-Separated Dynamic Mode Decomposition (SSDMD). The approach is shown to produce stable modes that can be used as a time-stepping model to predict the state of foF2 and hmF2 at a high time resolution. We demonstrate the SSDMD method on data sets covering periods of high and low solar activity as well as low, mid, and high latitude locations.


Connected comp. 1
Connected comp. 2
Connected comp. 3
Connected comp. 4
Connected comp. 5
Connected comp. 6
Connected comp. 7
Height (km)
Time (hours)

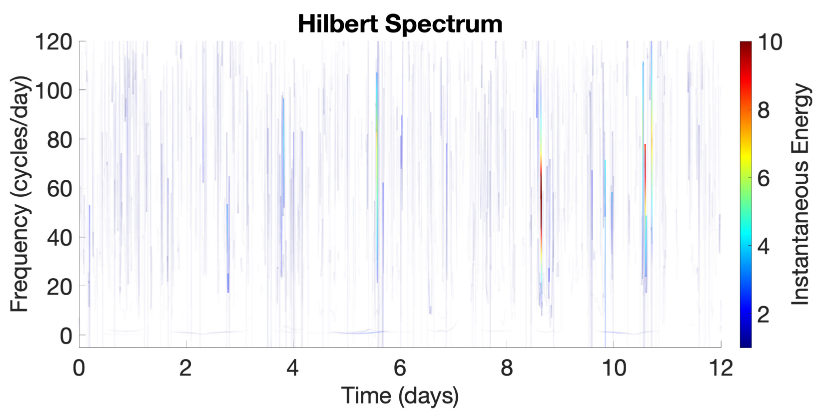

Connected comp. 1
Connected comp. 2
Connected comp. 3
Connected comp. 4
Connected comp. 5
Connected comp. 6
Connected comp. 7
Height (km)
Time (days)
Freq. (MHz)



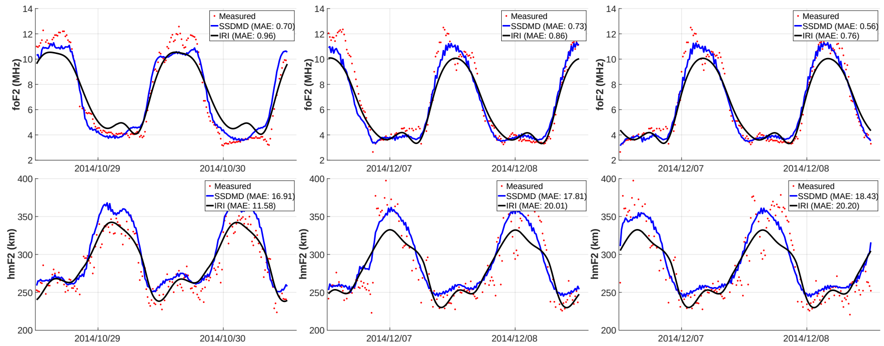
Correlation matrix
Wavelet index
Wavelet index
Correlation Graph



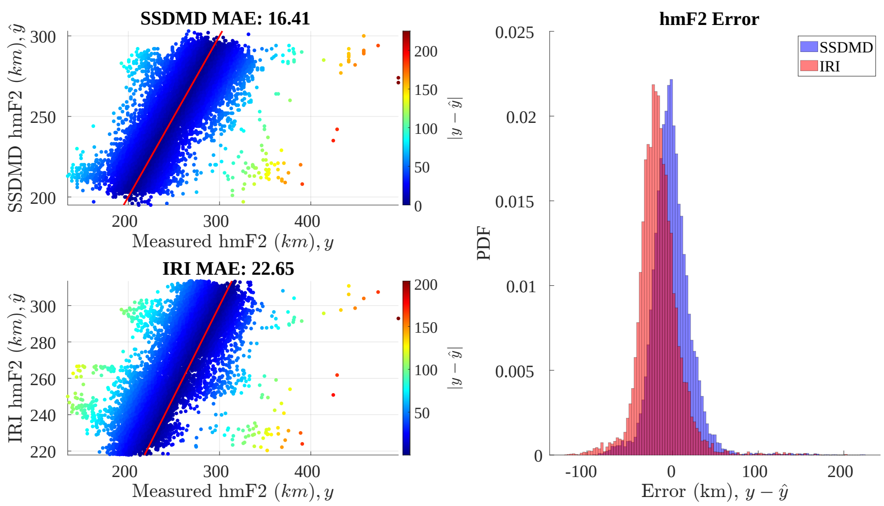
Boulder, CO (2019/10/05 - 2019/10/17)
Height (km)
Time (days)
Plasma Frequency (MHz)



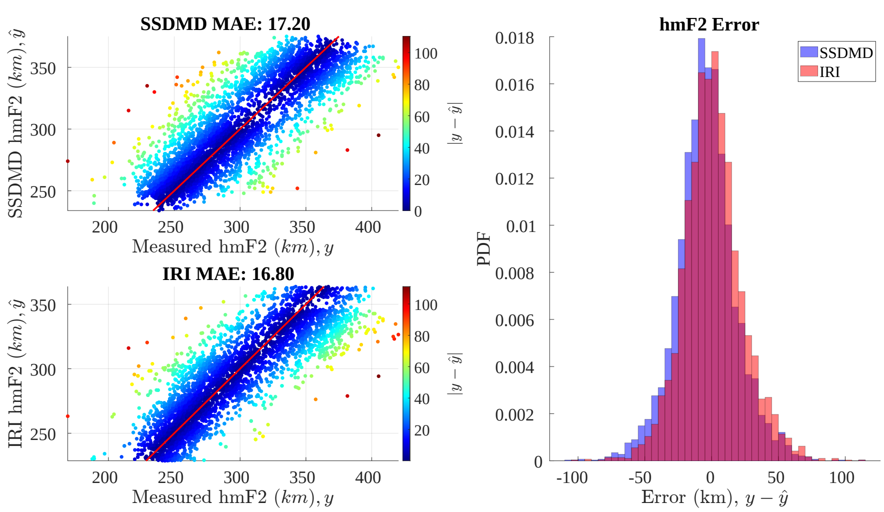
Hilbert Spectrum
Frequency (cycles/day)
Time (days)
Instantaneous Energy



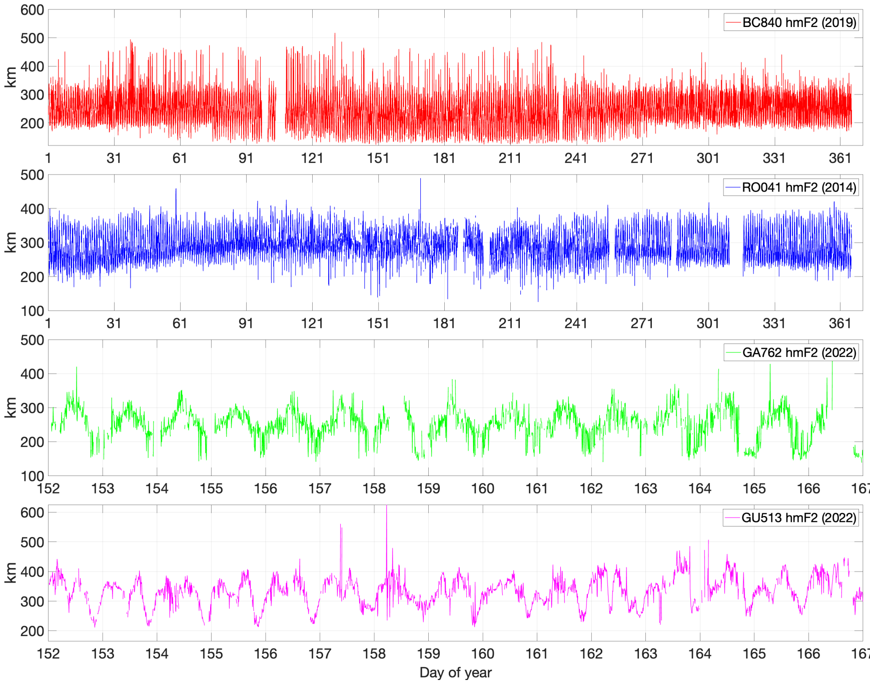
Measured
SSDMD (MAE: 0.70)
IRI (MAE: 0.96)
SSDMD (MAE: 0.73)
IRI (MAE: 0.86)
SSDMD (MAE: 0.56)
IRI (MAE: 0.76)
foF2 (MHz)
SSDMD (MAE: 16.91)
IRI (MAE: 11.58)
SSDMD (MAE: 17.81)
IRI (MAE: 20.01)
SSDMD (MAE: 18.43)
IRI (MAE: 20.20)
hmF2 (km)
2014/10/29
2014/10/30
2014/12/07
2014/12/08

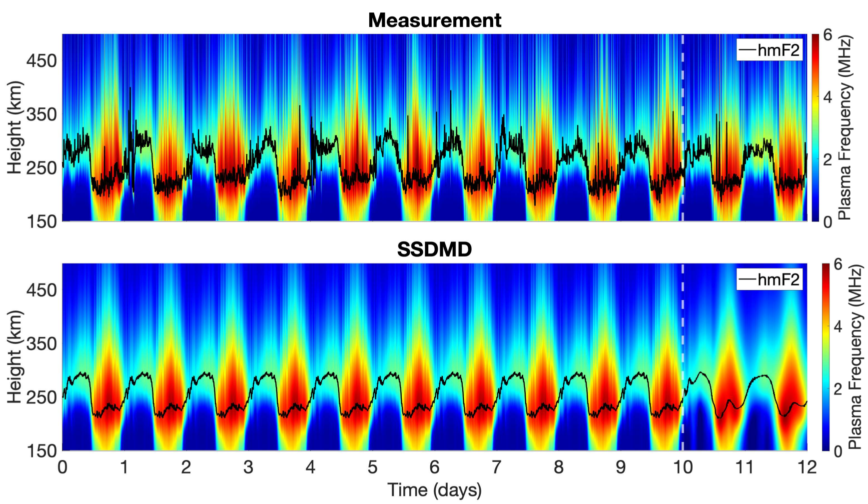


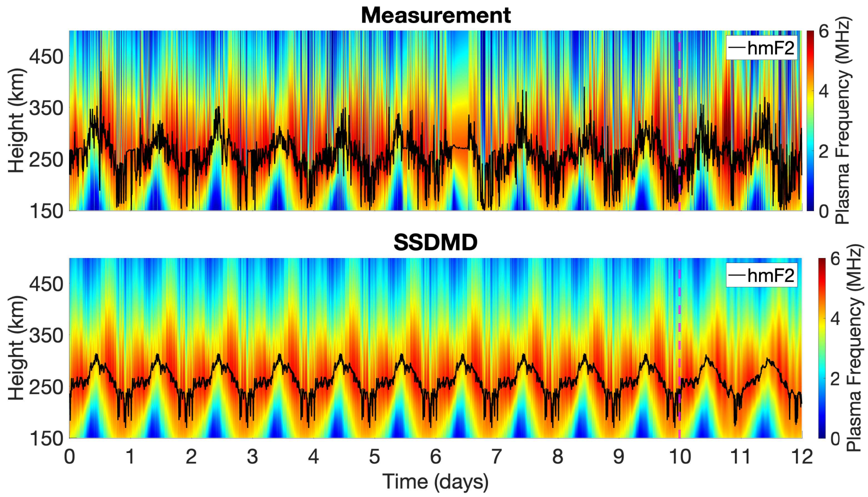
SSDMD MAE: 16.41
SSDMD hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
IRI MAE: 22.65
IRI hmF2 $(km), \hat{y}$
$|y - \hat{y}|$
hmF2 Error
SSDMD
IRI
PDF
Error (km), $y - \hat{y}$

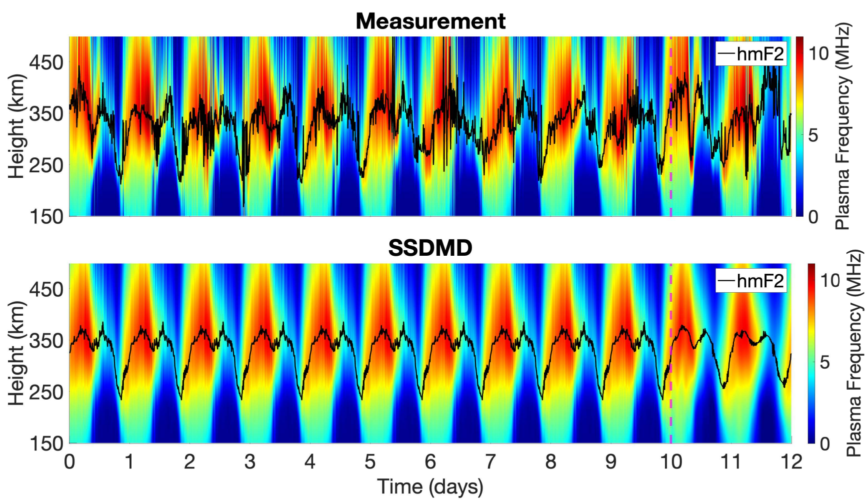


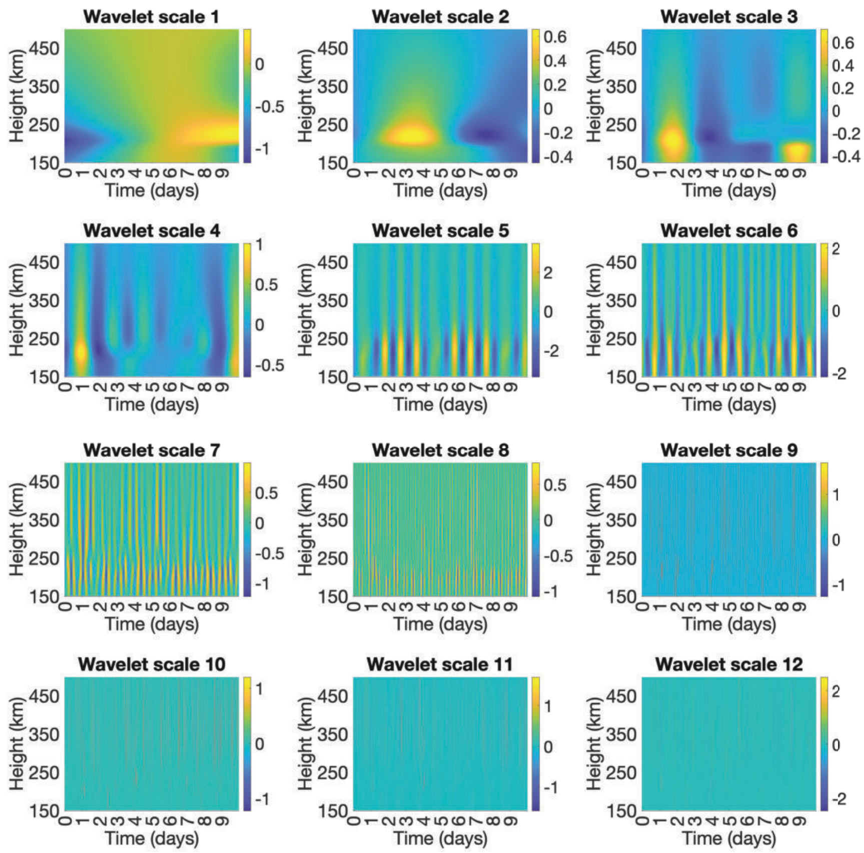
SSDMD MAE: 17.20
SSDMD hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
IRI MAE: 16.80
IRI hmF2 $(km), \hat{y}$
$|y - \hat{y}|$
hmF2 Error
SSDMD
IRI
PDF
Error (km), $y - \hat{y}$


BC840 hmF2 (2019)
RO041 hmF2 (2014)
GA762 hmF2 (2022)
GU513 hmF2 (2022)
km
Day of year


Measurement
SSDMD
hmF2
Height (km)
Plasma Frequency (MHz)
Time (days)


Measurement
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
0
2
4
6
SSDMD
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
0
2
4
6
0 1 2 3 4 5 6 7 8 9 10 11 12
Time (days)


Measurement
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
0
5
10
SSDMD
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
0
5
10
0 1 2 3 4 5 6 7 8 9 10 11 12
Time (days)


Wavelet scale 1
Wavelet scale 2
Wavelet scale 3
Wavelet scale 4
Wavelet scale 5
Wavelet scale 6
Wavelet scale 7
Wavelet scale 8
Wavelet scale 9
Wavelet scale 10
Wavelet scale 11
Wavelet scale 12
Height (km)
150
250
350
450
Time (days)
0 1 2 3 4 5 6 7 8 9

Scale-Separated Dynamic Mode Decomposition and Ionospheric Forecasting

Daniel J. Alford-Lago[1,2,3], Christopher W. Curtis[2], Alexander T. Ihler[3], Katherine A. Zawdie[4]

[1]Atmospheric Propagation Branch, Naval Information Warfare Center Pacific, San Diego, California, USA
[2]Department of Mathematics and Statistics, San Diego State University, San Diego, California, USA
[3]Department of Computer Science, UC Irvine, Irvine, California, USA
[4]Space Science Division, Naval Research Laboratory, Washington, District of Columbia, USA

Key Points:

- We present a method to adapt the Dynamic Mode Decomposition algorithm to work on a time series of ionospheric sounder profiles
- The method accounts for multiscale fluctuations in the time series using wavelet decompositions and builds a dynamical model from data alone
- The method can be used to forecast the foF2 and hmF2 parameters in near-real time using relatively short measurements from a sounder

Corresponding author: Daniel J. Alford-Lago, daniel.j.alford-lago.civ@us.navy.mil

Abstract

We present a method for forecasting the foF2 and hmF2 parameters using modal decompositions from measured ionospheric electron density profiles. Our method is based on Dynamic Mode Decomposition (DMD), which provides a means of determining spatiotemporal modes from measurements alone. Our proposed extensions to DMD use wavelet decompositions that provide separation of a wide range of high-intensity, transient temporal scales in the measured data. This scale separation allows for DMD models to be fit on each scale individually, and we show that together they generate a more accurate forecast of the time-evolution of the F-layer peak. We call this method the Scale-Separated Dynamic Mode Decomposition (SSDMD). The approach is shown to produce stable modes that can be used as a time-stepping model to predict the state of foF2 and hmF2 at a high time resolution. We demonstrate the SSDMD method on data sets covering periods of high and low solar activity as well as low, mid, and high latitude locations.

Plain Language Summary

Understanding the current and future state of Earth's ionosphere plays an essential role in many global communications and radar applications. However, generating accurate forecasts of it is challenging due to the complex physics that drive the dynamics. Additionally, measurements of the ionosphere show that there is a wide frequency range of fluctuations that occur in those measurements. We overcome both the complexity of the physics and the multiscale phenomena by applying methods from signals processing and machine learning to separate the various time scales over which these fluctuations arise. However, we do this in such a way that preserves strong couplings between the scales. We then demonstrate how to construct a forecast model from these separated scales. This approach to ionospheric forecasting is both equation-free and data-driven, and it is shown to have a modest improvement in accuracy over the current state-of-the-art.

1 Introduction

The need for accurate modeling and forecasting of the prevailing space weather conditions continues to play a critical role in the development and operation of a variety of radio communications and radar applications. The Earth's ionosphere is of particular interest as it provides a medium for the propagation of radio waves far beyond the horizon (Ratcliffe, 1959; Budden, 1985; Davies, 1990). As a result, the ionosphere has been the subject of intense study for decades, and efforts to enhance our ability to model and predict the vertical plasma density profile continue to this day. Parameterizations of the height-dependent structure of the ionosphere include specifying the maximum plasma density value and the height at which it occurs. This peak in the plasma density profile is known as the F2-layer critical frequency, *foF2*, and is generally given in units of megahertz (MHz). The altitude at which the foF2 occurs is called *hmF2* and has units of kilometers (km). Together, these two parameters specify a crucial point in the local ionosphere that can have a considerable impact on radio propagation. Specifically, foF2 and hmF2 will affect the reflection height and thus ground distance that a radio wave at a given frequency will reach (Fagre et al., 2019). Therefore, misrepresenting the peak of the plasma density profile has immediate implications for military, commercial, and civilian applications. In general, there are two modeling approaches for ionospheric specification: physics-based and empirical.

In physics-based models, the equations of fluid mechanics and magnetohydrodynamics are solved. However, the ionosphere is driven by many exogenous systems, including solar and geomagnetic activity, tidal forcing from the lower troposphere (H. L. Liu, 2016), and thermospheric general circulation (Killeen, 1987). This means that while the physics are relatively well-understood, careful specification of these drivers is required in order to produce accurate simulations and forecasts. Additionally, even when physics-based models such as the thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (TIME-GCM) (Dickinson et al., 1981; Roble & Ridley, 1994; Roble, 1995) and SAMI3 (Huba et al., 2000; Huba & Krall, 2013) offer accurate modeling capability, they often underestimate the variance observed in the measurements of the ionospheric plasma density (Zawdie et al., 2020).

On the other hand, empirical models, such as the International Reference Ionosphere (IRI), are generally less intensive to run but require large quantities of data from many different sources to account for the complex interactions between the various space weather systems. These sources include estimates from Mass Spectrometer Incoherent Scatter Radar (MSIS) to provide neutral composition derived from years of ground and space-based observations (Picone et al., 2002), as well as vertical soundings for the bottom-side, GPS-based observations of the total electron content (TEC), and *in situ* satellite measurements for the relevant ion species composition (Bilitza, 2001). Such an undertaking requires decades of dedicated service with international collaboration and has resulted in IRI becoming the official International Standardization Organization (ISO) standard for the ionosphere. Nevertheless, IRI provides only statistical estimates of the monthly average plasma density given several user-defined inputs such as solar activity via the monthly smoothed sunspot number and geomagnetic activity rather than simulating the dynamics.

More recently, determining reduced-order models (ROM) from data has been explored. In (Mehta et al., 2018), a quasi-physical dynamic ROM is obtained for the thermospheric mass density using the thermosphere-ionosphere-electrodynamics general circulation model (TIE-GCM) (Richmond et al., 1992), a precursor to TIME-GCM, as the source of observations. This ROM is based on a modal decomposition technique known as Dynamic Mode Decomposition (DMD) in which a set of spatiotemporal modes are determined via a linear best fit to data snapshots of a dynamical system (Schmid, 2010; Mezić, 2005; Kutz et al., 2016). DMD has also been shown to be especially useful in many physics and engineering contexts, such as in (Curtis et al., 2019) where it was used to help identify structure in weakly turbulent flows. Prior work on adapting DMD to data with dynamics at multiple scales can be found in (Dylewsky et al., 2019; Kutz et al., 2015), and building DMD models for nonlinear systems using deep learning in (Alford-Lago et al., 2022).

Our approach is motivated by the prevalence of vertical ionospheric sounder stations worldwide. These sounders generate data streams at regular cadences regarding the height-dependent profile of the ionospheric plasma density. However, plasma irregularities and traveling ionospheric disturbances manifest as fluctuations in the electron density profile (EDP) and occur over a range of time scales. Furthermore, the spatial frequencies of these irregularities are shown to range from the atmospheric scale height, where fluctuations are driven by gravity, down to the ion gyroradius, where fluctuations are driven by Earth's magnetic field (Booker, 1979).

We therefore see that modal analysis and dimensional reduction techniques, which facilitate the identification of simpler features within relatively complex data, would be of great utility in the study and use of ionospheric data. Likewise, measurement driven modeling techniques which bypass the intricate physics modeling that has been necessary to date to develop predictive capabilities would be especially desirable. To this end, we propose nontrivial extensions of DMD by way of wavelet decompositions that separate scales in a time series of EDPs. We call this method Scale-Separated DMD (SS-DMD) and demonstrate its utility in obtaining a dynamic model of the local ionospheric peak density from a relatively short recording of data.

SSDMD provides a novel approach to predicting the parameters foF2 and hmF2 that does not model their time evolution directly but instead uses the entire EDP time series to build a high-dimensional, expressive model for the dynamics. Our key contribution is incorporating a wavelet decomposition step and correlation analysis before applying DMD to the data. We find that critical couplings between scales that impact the stable evolution of DMD modes are preserved by grouping certain scales back together. These groupings are based on a one-step correlation that relates to how DMD is optimized. We find that the complete EDP forecasts from the method produce reasonable results in the F-region. However, the true utility of the method is the accuracy with which it predicts the foF2 and hmF2 parameters.

IRI was chosen for model comparison in this study because it is recognized as the official standard for the ionosphere by ISO, the International Union of Radio Science (URSI), the Committee on Space Research (COSPAR), and the European Cooperation for Space Standardization (ECCS) (Bilitza, 2018). While the number of ionospheric forecasting models seems to grow each year, we chose to use IRI as the gold standard because of its wide use in the community, (Bilitza, 2001) having over 1,000 citations at the writing of this paper, and is accessible to the research community through simple programming APIs. While there are variants of IRI that employ more sophisticated techniques such as assimilation of real-time data (Galkin et al., 2012), these models are more complex and generally less accessible to the public. Moreover, the goal of this paper and the SSDMD model itself is not to outperform the most advanced, high-fidelity ionospheric models. Instead, we aim to provide a simple approach to forecasting key parameters using minimal amounts of data while providing reasonably accurate results that are on par with the most common and established methods.

Note that many existing ionospheric forecast models require the specification of solar and geomagnetic drivers, often through the sunspot number and the station K and A indices. In this paper, however, we will show that a short-term forecast of the foF2 and hmF2 for a single-station sounder is indeed obtainable purely through modeling the variations observed during a 10-day period. While other attempts at forecasting ionospheric parameters without specifying drivers or control variables have seen success, see (Wang et al., 2020; Grzesiak et al., 2018; Stanislawska & Zbyszynski, 2001), we show that straightforward scale separation enables the use of powerful data-driven methods such as DMD. Of course, such an approach will not capture storms or large perturbations to the EDP that one would see with the appropriate exogenous control variables. Nevertheless, it lifts the burden of also having to forecast the drivers themselves and instead provides a lightweight, real-time method of forecasting the foF2 and hmF2. Additionally, observations of solar and geomagnetic activity are not widely available at the time resolution our method is set up to model. Many ionospheric sounding systems can produce measurements at a cadence of 5-minutes, whereas sunspot number and station indices are only available as averages over several hours.

This paper will provide the necessary background and algorithmic details to perform SSDMD on a time series of EDPs, and is organized as follows. In Section 2.1, we present the DMD algorithm to compute spatial modes with time-evolving dynamics. Then, in Section 2.2, we demonstrate how we generate a scale-separated expansion of a signal using wavelet decompositions. Sections 2.3 and 2.4 then describe how we determine strong couplings across scales in the time series and average across them to produce an SSDMD model. Finally, Section 3 presents our results from this analysis on measured data from several Digisonde vertical sounders (Reinisch & Galkin, 2011).

2 Method

The SSDMD method presented here will generate a near-term, e.g., 48-hour, forecast of the local ionospheric conditions using a time series of EDPs from a vertical incidence sounder. In particular, we will use this model to generate a forecast of the peak plasma density, foF2, and height, hmF2. The method consists of four primary steps:

1. Use 1-dimensional wavelet decompositions at each fixed height in the data to separate fluctuations at different time scales and reconstruct the signal with each scale individually.
2. Compute one-step correlations across each scale reconstruction, determine which scales are strongly correlated, and add them together to form *connected components*.
3. Average each connected component over 24-hour lags.
4. Perform DMD on the averaged connected components to obtain a set of modes and eigenvalues for each.

This algorithm will result in a separate DMD model for each of the averaged connected components. However, all these models will sum coherently to form a final reconstruction of the profile time series and predictions of its future state. From the forecasted profiles, we then compute the foF2 and hmF2 parameters.

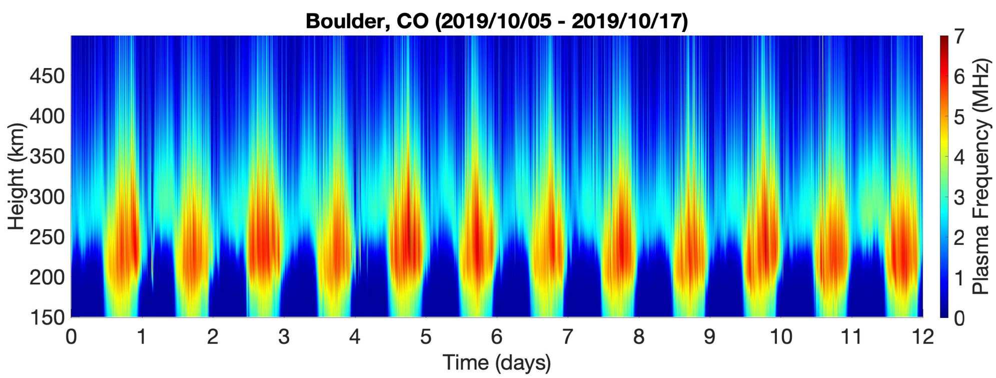


Figure 1: Dataset 1, a profilogram from the Digisonde Boulder, CO station covering the days of October 05, 2019 to October 17, 2019. Profiles were measured every 5 minutes.

The data used in this study are time series of ionospheric EDPs and their respective foF2 and hmF2 parameters gathered from two repositories, the Lowell GIRO Data Center digital ionogram database (Didbase) and the NOAA National Centers for Environmental Information (NCEI) Mirrion 2 data mirror. We will use a 12-day snippet, called Dataset 1, from a station in Boulder, Colorado, covering the dates 2019/October/05 to 2019/October/17 to illustrate each of the four steps of the SSDMD method above. This period of observation occurred near the last solar minimum yet still exhibits a wide spectrum of oscillations in the profile.

Figure 1 shows Dataset 1 as a profilogram, which we have preprocessed by interpolating the raw sounder profiles to a regular 1km resolution height grid and then clipped below 150km. This is done because our model is intended to capture the dynamics of the F-layer parameters of the ionosphere. The following sections will now illustrate each step in SSDMD, starting with a description of the DMD method since it forms the basis of SSDMD.

2.1 Dynamic Mode Decomposition

DMD provides a method of finding a one-step, linear best-fit transformation from a time series of data that maps any observation in the series one time-step into the future. We start with a series of measurements of the system

$$\mathbf{Y} = \{\mathbf{y}_1 \ \mathbf{y}_2 \cdots \mathbf{y}_{N_\mathrm{T}}\}\,, \tag{1}$$

where $\mathbf{y}_k = \mathbf{y}(t_k) \in \mathbb{R}^{N_S}$ is a snapshot of the system at time t_k, thus $\mathbf{Y} \in \mathbb{R}^{N_S \times N_T}$. In the case of Dataset 1, each snapshot is a measurement of the vertical profile so each column in $\mathbf{Y}$ is an EDP. We assume a regular measurement cadence with $t_k = k\delta t$ for some time step δt, though in general this is not a requirement. From this, we create two new matrices

$$\mathbf{Y}_- = \{\mathbf{y}_1 \ \mathbf{y}_2 \cdots \mathbf{y}_{N_T-1}\} \quad \text{and} \quad \mathbf{Y}_+ = \{\mathbf{y}_2 \ \mathbf{y}_3 \cdots \mathbf{y}_{N_T}\} \tag{2}$$

and find a matrix $\mathbf{K} \in \mathbb{R}^{N_S \times N_S}$ such that

$$\mathbf{K}\mathbf{Y}_- = \mathbf{Y}_+. \tag{3}$$

This can be done simply via regression by solving the following optimization problem,

$$\mathbf{K}_o = \underset{\mathbf{K}}{\operatorname{argmin}} \, ||\mathbf{Y}_+ - \mathbf{K}\mathbf{Y}_-||_F^2 = \mathbf{Y}_+\mathbf{Y}_-^\dagger, \tag{4}$$

where $||\cdot||_F$ denotes the Frobenius norm and $\mathbf{Y}_-^\dagger$ denotes the Moore-Penrose inverse of $\mathbf{Y}_-$. The DMD model is then given by the eigendecomposition of the matrix $\mathbf{K}_o$, however, solving (4) directly can generate highly unstable results due to ill-conditioning in $\mathbf{Y}_-$. To address this, it is common in the DMD literature to use the singular-value decomposition (SVD) of $\mathbf{Y}_-$ and apply a threshold to keep only the most significant singular values. If the SVD of $\mathbf{Y}_-$ is

$$\mathbf{Y}_- = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*, \tag{5}$$

then introducing a threshold, $c_{\text{svd}} > 0$, we truncate the columns of $\mathbf{U}$ and $\mathbf{V}$ corresponding to the singular values, Σ_{jj}, such that

$$\log_{10}\left(\frac{\Sigma_{jj}}{\Sigma_{11}}\right) > -c_{\text{svd}}, \tag{6}$$

where Σ_{jj} are entries along the diagonal of $\mathbf{\Sigma}$ and are ordered such that

$$\Sigma_{11} \geq \Sigma_{22} \geq \cdots \geq \Sigma_{N_S N_S}. \tag{7}$$

We label the truncated versions of $\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}$ as $\tilde{\mathbf{U}}$, $\tilde{\mathbf{\Sigma}}$, and $\tilde{\mathbf{V}}$ respectively. A straightforward approximation of Equation (4) can then be given by

$$\mathbf{K}_o \approx \mathbf{Y}_+\tilde{\mathbf{V}}\tilde{\mathbf{\Sigma}}^{-1}\tilde{\mathbf{U}}^*. \tag{8}$$

Note, $\mathbf{K}_o$ will be an $N_S \times N_S$ matrix, so when N_S is very large it may be computationally expensive to compute the eigendecomposition; see (Tu et al., 2014) for alternate formulations of DMD when this is the case. However, we found that the EDP data from a single sounding station is not high-dimensional enough to require these alternate forms. Instead, we simply compute the DMD modes and eigenvalues of $\mathbf{K}_o$ through the diagonalization

$$\tilde{\mathbf{K}}_o = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^{-1}, \tag{9}$$

where $\mathbf{W}$ is a matrix whose columns are eigenvectors, or DMD modes, and $\mathbf{\Lambda}$ is a diagonal matrix of DMD eigenvalues. For a given δt representing the amount of time which has passed from observation $\mathbf{y}_k$ to $\mathbf{y}_{k+1}$, we construct a continuous-time model of the system,

$$\mathbf{y}(t) \approx \mathbf{W}\mathbf{\Lambda}^{t/\delta t}\mathbf{W}^\dagger\mathbf{y}(0), \tag{10}$$

where $\mathbf{y}(0)$ is some initial condition. Note that this decomposition provides a time stepping mechanism for reconstructing our time series that we may use for forecasting.

Comparisons of DMD to the well-established Empirical Orthogonal Function (EOF) analysis may be drawn. In practice, EOF models use Principal Component Analysis (PCA) to decompose the data into linear combinations of orthogonal functions. Fourier expansions of modulating coefficients for each component then provide variation over monthly and solar cycle scales; see (C. Liu et al., 2008; Zhang et al., 2009, 2014; Mehta & Linares, 2017; Li et al., 2021) for in-depth description of EOF analysis for space weather. This has the advantage of including proxies for external drivers such as the F10.7-cm solar flux in the forecast. Nevertheless, such indices are not readily available on the time scales that we are able to measure ionospheric profiles and provide little additional input for a 24- to 48-hour forecast. Furthermore, EOF models are restricted to an orthogonal basis of functions for the dynamics due to the use of PCA. The DMD modes have no such restriction since they are derived from the eigendecomposition of the $\mathbf{K}_o$ matrix. Another major difference between our method and conventional EOF models is we separate the various time scales in the data prior to fitting the DMD modes and eigenvalues.

Thus, beyond just producing a modal decomposition from data, the DMD method gives a time-evolving model for said data through the spectra of the $\mathbf{K}$ matrix. Further connections between DMD and dynamical systems analysis can be established through its relationship with the Koopman operator (Koopman, 1931); see Appendix A. While a generally successful approach, this straightforward implementation struggles with multiscale data or any data that has both very small and very large gradients from snapshot to snapshot due to the one-step regression in Equation 4. This motivates the use of some form of temporal scale separation.

2.2 Scale Separation of EDP Time Series

The primary contribution of this paper is to provide a method of adapting the DMD algorithm to work on data with fluctuations at multiple scales, as is the case when modeling EDP measurements. The need to account for these oscillations is motivated by the Hilbert spectrum of a slice through Dataset 1 at a vertical height of 400km. At this altitude, we see there is a significant degree of instantaneous energy at frequencies much higher than diurnal variation (1 cycle/day); see Figure 2. These relatively high-frequency, transient events complicate direct applications of DMD, but do not necessarily represent noise that should be filtered out.

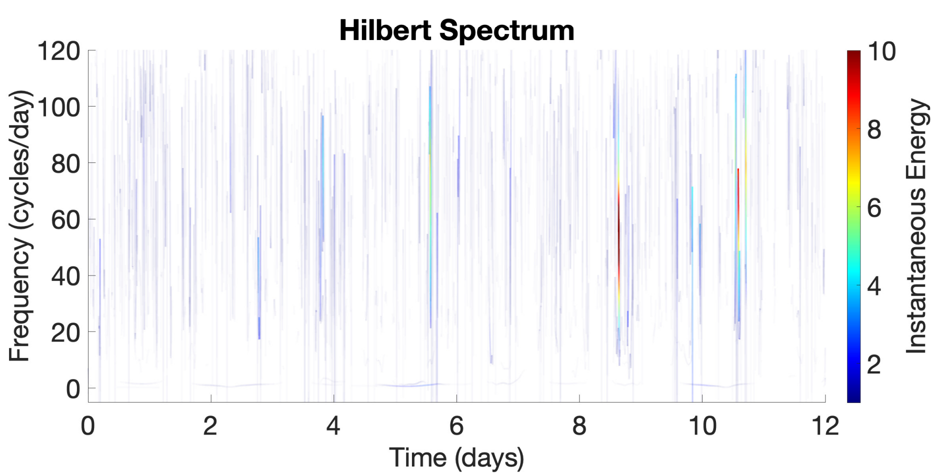


Figure 2: The affiliated Hilbert spectrum for a slice through Dataset 1 at a height of 400km. The Hilbert spectrum plot reveals the instantaneous energy in the data as a function of time and frequency. The stable diurnal oscillation can be see near 1 cycle/day, while various time localized, spurious oscillations occur throughout at frequencies that are an order of magnitude higher.

We therefore use a multiresolution analysis by way of 1-dimensional wavelet decompositions to facilitate DMD; see (Mallat, 2009, 1989) for in-depth theory and applications of wavelet decompositions. For a given time series $\mathbf{y}(t) \in \mathbb{R}^{N_s}$ representing vector observations of EDPs, we decompose each height in the time series into N_{lvl} levels, such that

$$\mathbf{y}(t) \approx \sum_{j=1}^{N_{lvl}+1} \mathbf{d}_j(t), \tag{11}$$

where $\mathbf{d}_j(t) \in \mathbb{R}^{N_s}$, such that

$$\mathbf{d}_j(t) = \sum_{n=-M_f}^{M_f} \mathbf{d}_{j,n}\psi_{j,n}(t),\ 1 \le j \le N_{lvl}, \tag{12}$$

and

$$\mathbf{d}_{N_{lvl}+1}(t) = \sum_{n=-M_f}^{M_f} \mathbf{d}_{N_{lvl}+1,n}\phi_{N_{lvl},n}(t), \tag{13}$$

where $\psi(t)$ and $\phi(t)$ are the wavelet and scaling functions of the decomposition, respectively,

$$\begin{aligned} \psi_{j,n}(t) &= \sqrt{2}^{-j}\psi\left(2^{-j}t - n\right), \\ \phi_{N_{lvl},n}(t) &= \sqrt{2}^{-N_{lvl}}\phi\left(2^{-N_{lvl}}t - n\right). \end{aligned} \tag{14}$$

The vectors $\mathbf{d}_{j,n}$, $1 \le j \le N_{lvl}$, denote the *detail coefficients* at the j^{th} scale while $\mathbf{d}_{N_{lvl}+1,n}$ denotes the *approximation coefficients* at the terminal scale.

With the wavelet decompositions performed independently at each height in the profile, the vector quantities $\mathbf{d}_j(t)$ represent only parts of the signal at the j^{th} scale at time t. Given our discrete time series from Equation (1), these vector quantities form the columns of a new set of data matrices,

$$\mathbf{Y}_j = \{\mathbf{d}_{j,1}\ \mathbf{d}_{j,2}\ \cdots\ \mathbf{d}_{j,N_T}\}, \tag{15}$$

which are reconstructions of the original data at each scale and sum coherently, so that $\mathbf{Y} = \sum_{j=1}^{N_{lvl}+1} \mathbf{Y}_j$.

In Figure 3, we have Dataset 1 expanded into 12 scale reconstructions. These scales further illustrate the multiscale nature of high-resolution EDP measurements, with fluctuations on the order of 1-2MHz in magnitude observed up to the fastest scales. These sub-diurnal oscillations can appear as broad-spectrum noise in the raw profilogram and can make modal decompositions like DMD quite challenging. Note that the diurnal oscillation itself does not appear until the 5^{th} or 6^{th} scale in Figure 3, and several longer-period trends are observed before the terminal scale. In the following section we will see how these oscillations can be highly correlated in terms of an optimal DMD one-step fit. Fourth-order Coiflets were used for the discrete wavelet transforms. The wavelet type is a model hyperparameter and may vary for different data sets. However, we found that this choice worked well for all test cases in this study.

2.3 Computing Correlations Across Scales

Applying DMD to each scale separately does not produce optimal results and can even produce DMD modes that are unstable and decay to zero or grow to infinity almost immediately. Instead, we found correlations across each of the scales can indicate strong dynamical couplings between them, and preserving these has a pronounced impact on the fidelity and stability of the DMD modes. Identifying the strength of these couplings required developing a measure of correlation that takes into account the role that the

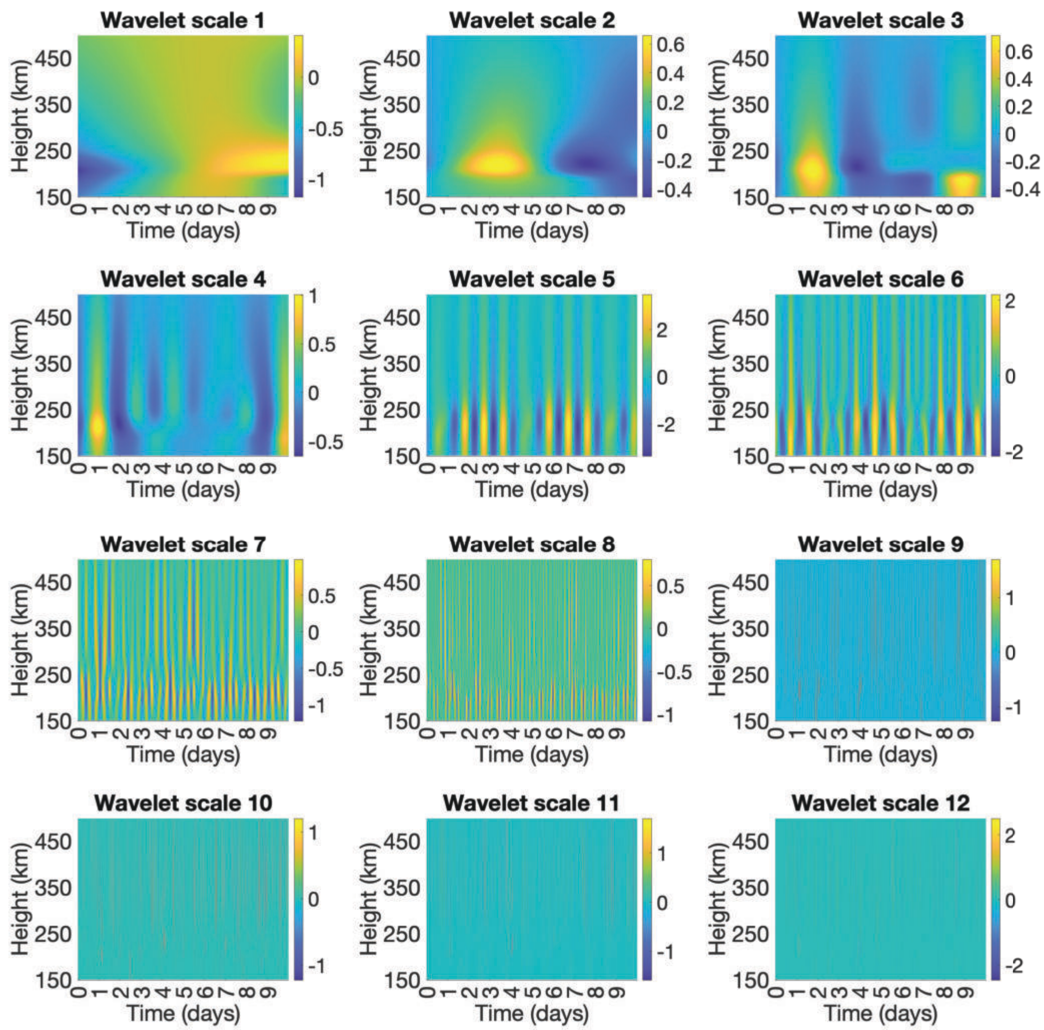


Figure 3: Dataset 1 decomposed into 12 scales. Each panel is a reconstruction of the full EDP time series using only the j^{th} scale coefficients from the wavelet decomposition at each height, $\mathbf{Y}_j$. The color axis represents plasma frequency in MHz.

matrix $\mathbf{K}_o$ plays in advancing the data forward in time. To this end, we defined the following correlation matrix $\mathbf{C}$ whose entries are given by

$$C_{jl} = \frac{1}{2}\left|\overline{\tilde{\mathbf{y}}_{j,+} \odot \tilde{\mathbf{y}}_{l,-}} + \overline{\tilde{\mathbf{y}}_{j,-} \odot \tilde{\mathbf{y}}_{l,+}}\right|, \tag{16}$$

with, $j, l \in 1, \ldots, N_{lvl} + 1$, and

$$\tilde{\mathbf{y}}_j = \frac{[\mathbf{Y}_j] - \overline{[\mathbf{Y}_j]}}{\left|\left|[\mathbf{Y}_j] - \overline{[\mathbf{Y}_j]}\right|\right|_{2,t}}. \tag{17}$$

The $\bar{\cdot}$ and $[\cdot]$ denote taking the mean in the time and space dimensions of the time series, respectively, $||\cdot||_{2,t}$ is an L_2-norm over time, and $\odot$ is the Hadamard product between two matrices. Finally, the $+$ and $-$ subscripts indicate shifting the time series forward or backward one time step as in Equation 2. Note that the full-dimensional EDP is reduced to an average for this correlation coefficient in Equation 17. This works because we have limited the time series to the upper F-region of the profile since we are concerned with forecasting the F-peak characteristics only.

Because $\mathbf{K}_o$ is optimized to advance any profile in the data one time step into the future, this correlation coefficient provides a quantitative means for comparing the time series across different timescales in the context of fitting optimal DMD modes. Then, by setting a threshold value, c_{corr}, we generate an adjacency matrix $\mathbf{A}$ with entries

$$A_{jl} = \begin{cases} 1, & |C_{jl}| \geq c_{\text{corr}} \\ 0, & |C_{jl}| < c_{\text{corr}} \end{cases} \tag{18}$$

The matrix $\mathbf{C}$ is symmetric, and so $\mathbf{A}$ is as well. Note, in practice these correlations will typically be larger for the longer time scales since we are looking at one-step correlations, with higher frequency oscillations becoming increasingly less correlated. The matrix $\mathbf{A}$ generates a graph G that indicates which of the $\mathbf{Y}_j$ scale reconstructions should be grouped back together to preserve their dynamic coupling.

Thus, for a given choice of threshold c_{corr}, we will have $N_C \leq N_{lvl}+1$ connected components within G. We then form N_C new time series by summing only the $\mathbf{Y}_j$ which belong to the same connected component,

$$\mathbf{Y}_n^{\text{C}} = \sum_{j \in G_n} \mathbf{Y}_j, \tag{19}$$

where $j \in G_n$ denotes the scales that are in the n^{th} connected component in G, and $\mathbf{Y}_n^{\text{C}}$ is the time series for the n^{th} connected component. Figure 4 shows the matrix $\mathbf{C}$ and the graph G for Dataset 1. Note that the first group consists of the bulk of the large scale features in the time series while the higher frequency scales remain on their own. However, this may not always be the case, and subgroups within the high frequency components could arise depending on the data observed.

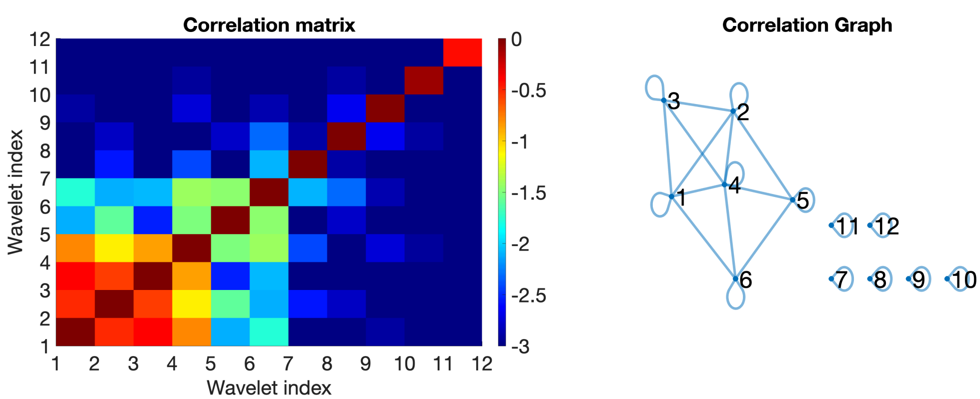


Figure 4: The correlation coefficient matrix $\mathbf{C}$ (left) and the corresponding graph G (right) indicating which scales are highly coupled. The correlation threshold $c_{corr} = -1.95$ was used for Dataset 1.

At this point, one could find a corresponding $\tilde{\mathbf{K}}_{o,n}$ via DMD and generate an affiliated expansion for each connected component so that the total time series can be approximated by

$$\mathbf{y}(t) \approx \sum_{n=1}^{N_C} \mathbf{W}_n \mathbf{\Lambda}_n^{t/\Delta t} \mathbf{W}_n^{\dagger} \mathbf{y}_{n,0}. \tag{20}$$

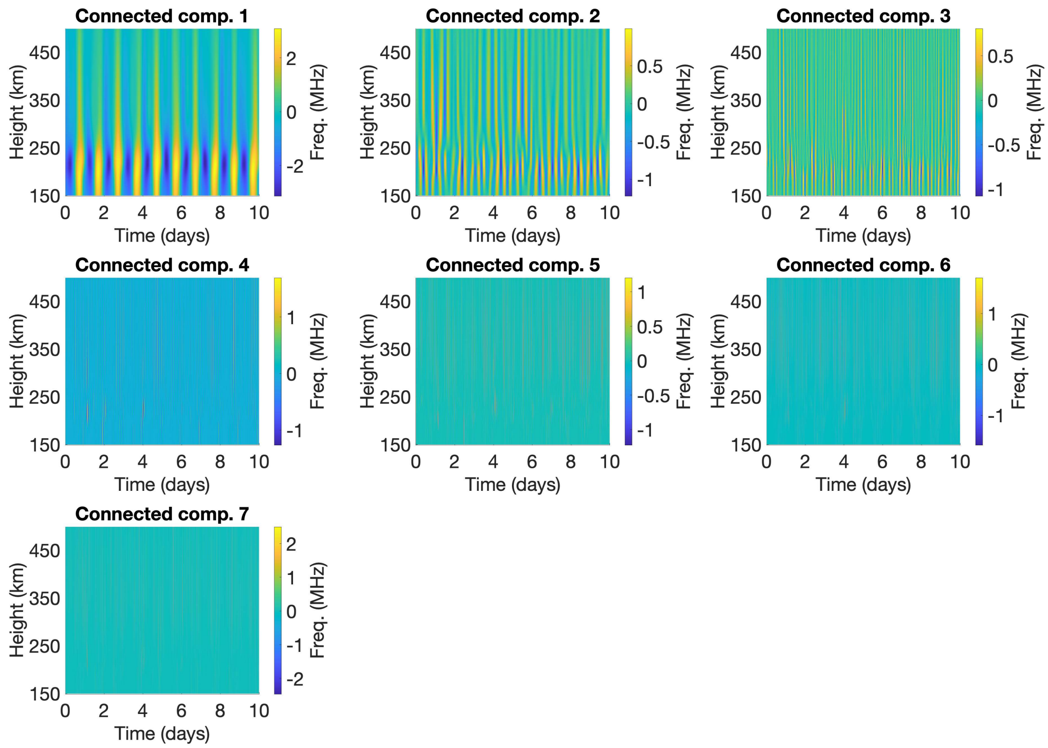


Figure 5: Dataset 1 decomposed into 7 connected components. Each component captures features of the data with strong correlations according to the one-step spatiotemporal coefficients.

However, we note that, using observations that span only several days in time, the EDP at a single sounding station is essentially memoryless after twenty-four hours have passed (Araujo-Pradere et al., 2005). This strongly suggests that before naively applying the DMD method to time series of arbitrary length, instead, we should first average the data across 24-hour cycles for the duration of our measurement period.

2.4 Averaging for DMD

Having decomposed the EDP time series into correlated time scales, we now have a collection of time series,

$$\mathbf{Y}_1^{\mathrm{C}},\ \mathbf{Y}_2^{\mathrm{C}},\ \cdots,\ \mathbf{Y}_{N_C}^{\mathrm{C}}, \tag{21}$$

that represent scales within the data set whose one-step correlations are relatively weak. We treat these as being essentially independent with respect to our DMD approximation.

Denoting the number of time steps in a full day as T_D and assuming that $N_T + 1$ is divisible by T_D, so that the data set represents the number of days N_D where

$$N_D = \frac{N_T + 1}{T_D}, \tag{22}$$

we isolate the mean signal over 24-hour cycles from the fluctuations about the mean for each $\mathbf{Y}_n^{\mathrm{C}}$. This creates two new affiliated time series for each connected component that have the properties,

$$\bar{\mathbf{y}}_n^{\mathrm{C}}(t_k + T_D) = \bar{\mathbf{y}}_n^{\mathrm{C}}(t_k), \tag{23}$$

and

$$\sum_{k=1}^{T_D} \hat{\mathbf{y}}_n^{\mathrm{C}}(t_k + mT_D) = 0, \quad m = 0, \ldots, N_D - 1, \tag{24}$$

where $\bar{\cdot}$ and $\hat{\cdot}$ denote the 24-hour mean signal and fluctuations about the 24-hour mean, respectively. The fluctuations in Equation 24 effectively represent the noise signal for each component. These may prove useful in future experiments to generate nonparametric error estimates, however, in this paper they are not used further since our goal is to forecast parameters derived from the profile. Taking the vector quantities, $\bar{\mathbf{y}}_n^{\mathrm{C}}$ to be columns of new mean-signal matrices we have

$$\bar{\mathbf{Y}}_n^{\mathrm{C}} = \left\{ \bar{\mathbf{y}}_{n,1}^{\mathrm{C}},\ \bar{\mathbf{y}}_{n,2}^{\mathrm{C}},\ \ldots,\ \bar{\mathbf{y}}_{n,T_D}^{\mathrm{C}} \right\}. \tag{25}$$

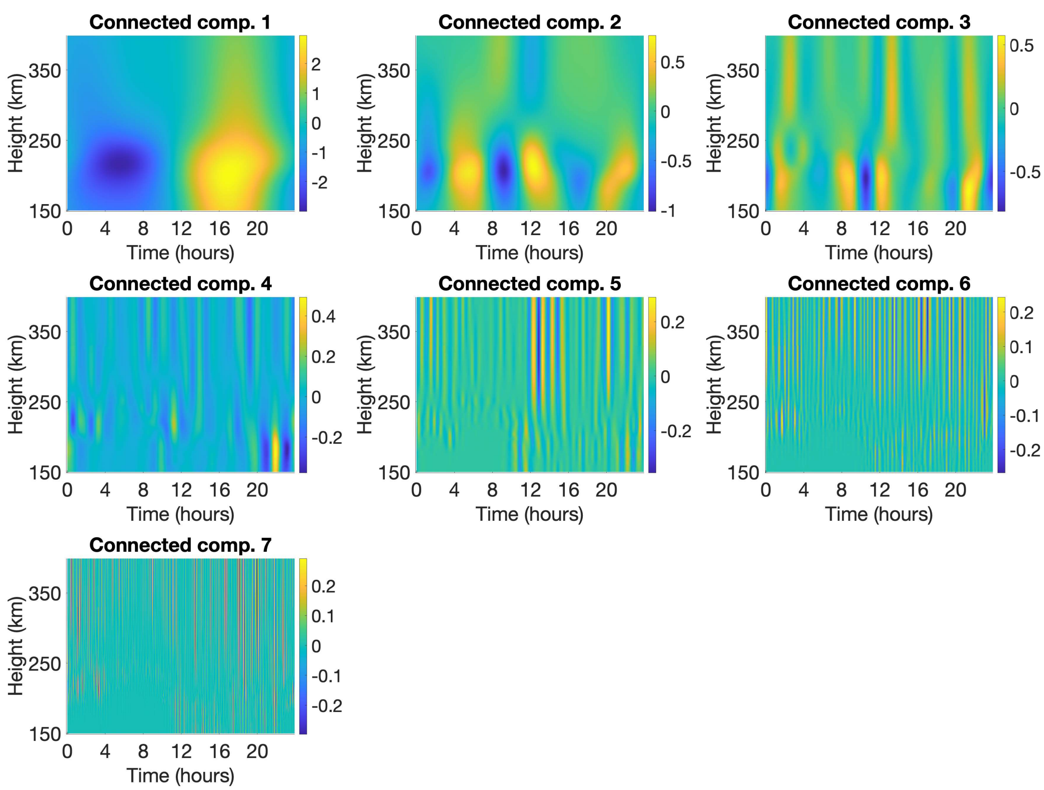


Figure 6: Dataset 1 connected components averaged over 24-hour lags, $\bar{\mathbf{Y}}_n^{C}$.

Figure 6 shows each $\bar{\mathbf{Y}}_n^{\mathrm{C}}$ for Dataset 1. These matrices represent the average plasma frequency oscillation over a given day at various scales in the dynamics. Therefore, this step acts as a denoising process that has minimal impact on the multiscale nature of the signal and reduces the amount of information that would be lost by simply filtering the raw EDP time series.

Finally, using Equation (10) on these 24-hour averaged and scale-correlated data, we generate a continuous-time DMD model for each connected component,

$$\bar{\mathbf{y}}_n^{\mathrm{C}}(t) \approx \mathbf{W}_n \mathbf{\Lambda}_n^{t/\Delta t} \mathbf{W}_n^{\dagger} \bar{\mathbf{y}}_{n,0}^{\mathrm{C}}. \tag{26}$$

Note that all of the N_C components sum coherently and form the final the SSDMD model,

$$\bar{\mathbf{y}}(t) \approx \sum_{n=1}^{N_C} \mathbf{W}_n \mathbf{\Lambda}_n^{t/\Delta t} \mathbf{W}_n^{\dagger} \bar{\mathbf{y}}_{n,0}^{\mathrm{C}}. \tag{27}$$

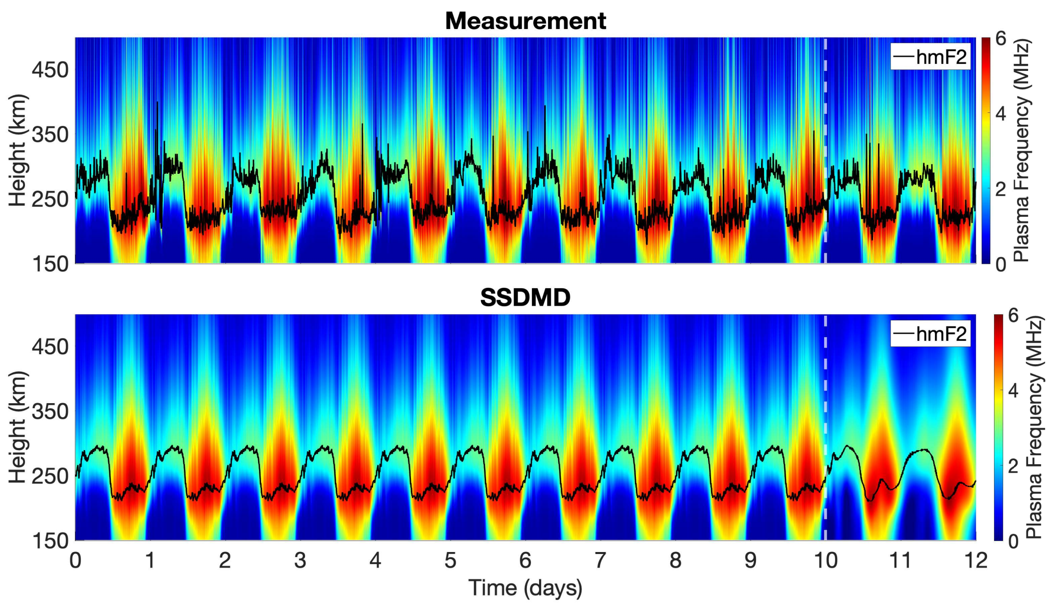


Figure 7: SSDMD reconstruction and forecast of Dataset 1. The vertical dotted white line denotes the transition from data used to fit the model to validation data. Black lines in each panel trace the hmF2 parameter.

Equation (27) is a model for the dynamics of the average that accounts for non-linear oscillations at multiple scales while preserving strong couplings between scales. See Appendix B for pseudocode of the complete SSDMD algorithm. Figure 7 depicts the result of this model applied to Dataset 1, using the first 10 days of data to generate the SSDMD model and then advancing the DMD modes via their eigenvalues out an additional 2 days as a forecast. The figure includes both the original measurement time series and the SSDMD reconstruction and forecast.

We compute the foF2 and hmF2 parameters by finding the peak frequency and height in the modeled EDPs. Figure 7 shows the predicted hmF2 and observed hmF2 overlayed on their respective EDP time series. The reconstruction of the first 10 days, i.e. the fitting data, appears excellent simply because it is advancing each profile a single time step. The remaining two days, however, illustrate the stability of the modes that have been determined through SSDMD, since we are iterating the DMD eigenvalues and using the last observed EDP from the training data as an initial condition. Thus, we have built a stable time-stepping model of foF2 and hmF2 using a dynamical model that utilizes the full EDP time series expanded over several time scales. In Section 3.2 we will explore the accuracy of the resultant foF2 and hmF2 forecasts in greater detail.

3 Results

3.1 Data Description

Data sets were gathered from Boulder, Colorado (40°N, -105.3°W) over 2019, and from Rome, Italy (41.9°N, 12.5°E) over 2014. The years 2019 and 2014 were roughly at the last solar minimum and solar maximum, respectively. These data sets will provide statistical estimates of how the proposed method performs at mid-latitudes during periods of high and low solar activity. Additionally, shorter data sets taken from Gakona, Alaska (62.38°N, 145°W) and Guam (13.62°N, 144.86°E) and will demonstrate the method's application in high-latitude and equatorial environments, respectively. Results presented for foF2 are in units of megahertz and hmF2 in kilometers unless otherwise labeled.

The sounder located in Boulder, Colorado (station name BC840) had a measurement cadence of 5 minutes in 2019, while the Rome, Italy sounder (station name RO041) measured profiles every 15 minutes in 2014. The shorter data sets from Gakona, Alaska (station name GA762) and Guam (station name GU513) both had cadences of 7.5 minutes. Table 1 summarizes the locations, times, and lengths of the data sets gathered for this study, and Figures 8 and 9 show time series of the foF2 and hmF2 parameters as measured at each station. Each data point in these time series has an affiliated EDP, but these are not shown for brevity. Missing values in the data are not used in the final error analysis.

All sounder stations generate estimates of the vertical EDP using the ARTIST5 algorithm to invert raw ionograms (Galkin & Reinisch, 2008). The EDP time series is limited to a height range of 150-500km. This is primarily because the plasma frequency in E-region at night dips low enough that it is outside the measurement bandwidth of the Digisonde sounders (Bibl et al., 1981). Because of this, the ARTIST5 inversion algorithm will generally output a default value, e.g., 0.2 MHz, in these regions for most of the nighttime profiles. These periods of constant plasma density complicate the fitting of an SSDMD model since they require inherently oscillatory modes to approximate a constant value. Above the peak plasma density, echoes from the sounder are no longer received, and a standard parameterized profile is fit to provide the topside plasma density. Thus, restricting the profiles to only the F-region helps ensure the SSDMD model is able to more accurately capture the dynamics of the F-layer parameters and minimizes the effects of these boundary regions.

	Boulder	**Rome**	**Gakona**	**Guam**
Station name	BC840	RO041	GA762	GU513
Year	2019	2014	2022	2022
Lat/Lon	40°N 105.3°W	41.9°N 12.5°E	62.38°N 145°W	13.62°N 144.86°E
Number of days	365	365	12	12
Measurement cadence	5 min.	15 min.	7.5 min.	7.5 min.
Solar cycle	min	max	mid	mid

Table 1: Summary of data gathered from Didbase sounder stations.

We used the IRI2016 model in Python with up-to-date solar and magnetic indices. IRI has many settings that allow the user to tweak parameters or turn certain submodels on or off. These settings are known as the *JF* switches. The version of IRI used in

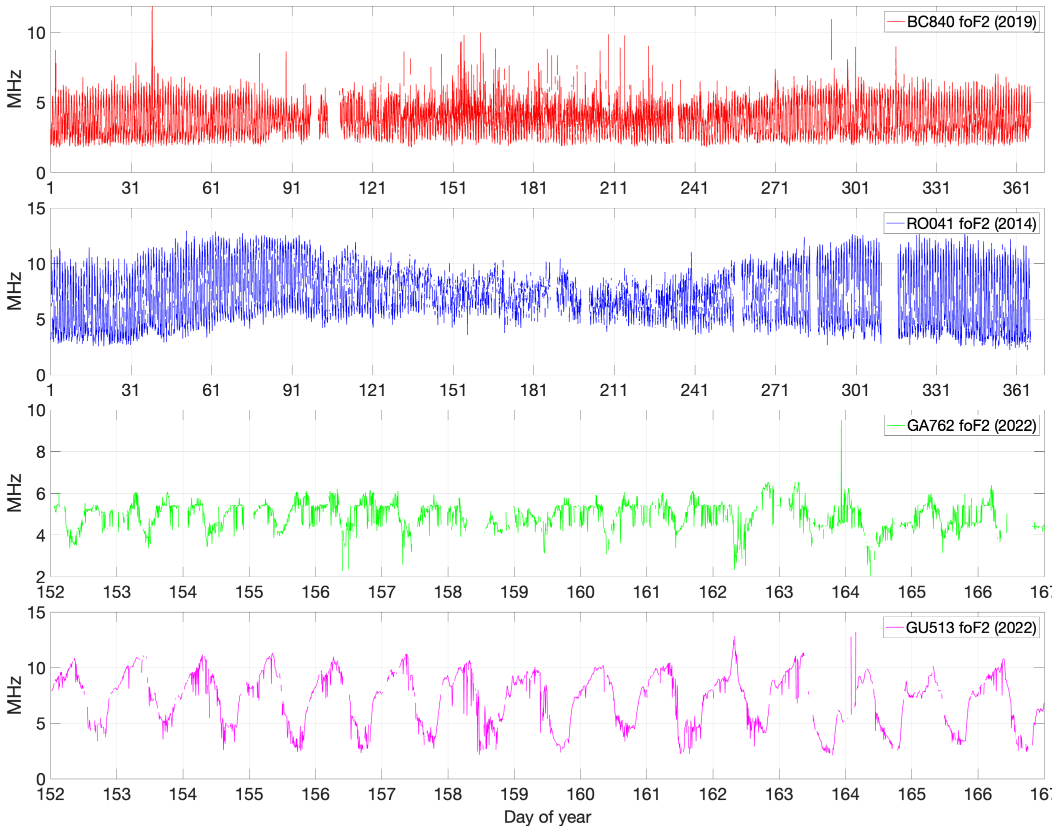


Figure 8: Time series of foF2 from the BC840 (red), RO041 (blue), GA762 (green), and GU513 (magenta) sounders. Note that the x-axis (day of year) has been zoomed in for the shorter data sets GA762 and GU513.

this paper had all the default JF values, which are found on the IRI model website. Time series of the EDP, foF2, and hmF2 were generated from IRI for each data set, and the EDPs were interpolated to the same vertical height grid as the sounder data.

There are several hyperparameters of the SSDMD model that must be set prior to fitting a model. The first is the correlation threshold from Equation 18 that determines how strongly scales must be correlated in order to form a connected component. This threshold currently requires manual tuning. We found a value of $c_{corr} = -1.95$ achieved good results for stations BC840, GA762, and GU513, while $c_{corr} = -1.75$ performed better for RO041. Generating more efficient ways of determining the optimal value for this parameter will be a topic of future research, though its value here was chosen such that the MAE of the foF2 and hmF2 parameters were minimized.

Another hyperparameter is the number of days used to fit the SSDMD model. Using long time series will result in more averaging over the 24-hour cycles, thus increasing bias in the forecast. We found that 10 days of EDPs worked reasonably well for all stations for short-term prediction. If one attempts a longer-term forecast, averaging over additional time lags may be necessary. The last hyperparameter of SSDMD is the threshold at which to truncate the singular values in the DMD step, Equation 6. This threshold was set to $c_{svd} = 6$, which worked well for all data sets. Lowering this threshold will result in fewer spectral pairs $(\lambda_j, \mathbf{w}_j)$ in the SSDMD model and thus reduces the number of modes used to generate the forecast. Table 2 summarizes these hyperparameters.

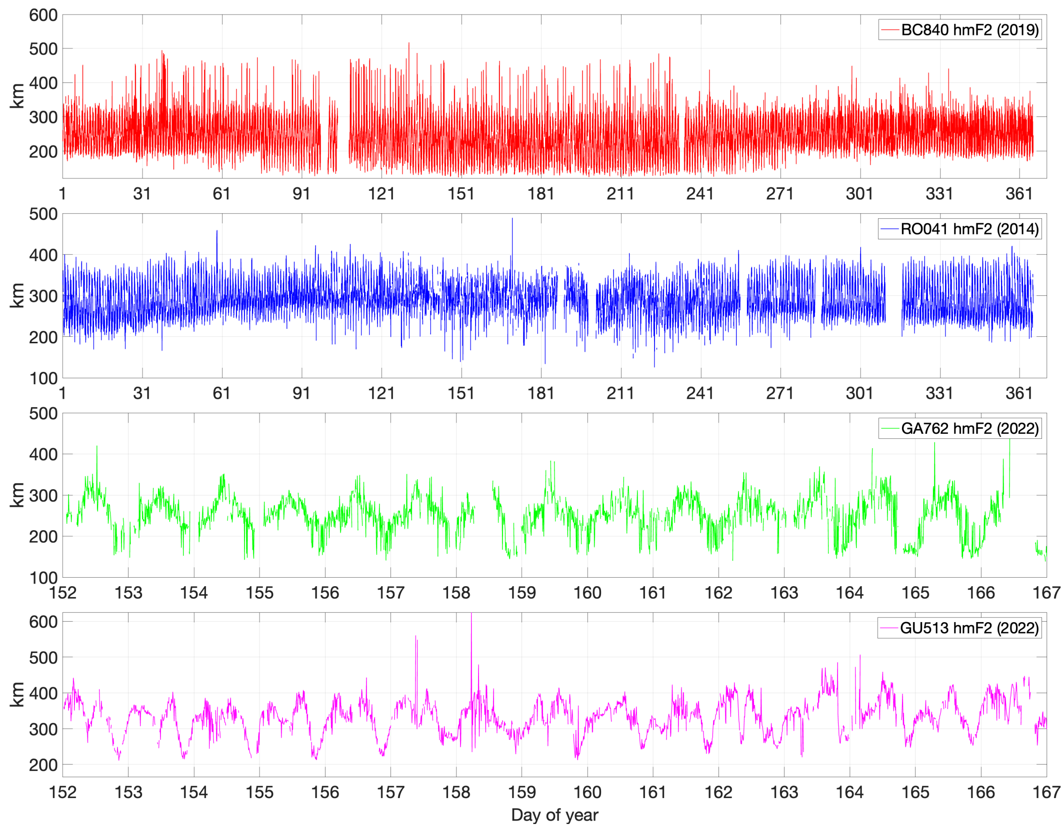


Figure 9: Time series of hmF2 from the BC840 (red), RO041 (blue), GA762 (green), and GU513 (magenta) sounders. Note that the x-axis (day of year) has been zoomed in for the shorter data sets GA762 and GU513.

SSDMD Parameter	**Value**
Num. days for fit	10
Num. days forecast	2
c_{corr}	-1.95 (BC840, GA762, GU513) / -1.75 (RO041)
c_{svd}	6
Wavelet type	coiflet 4^{th} order

Table 2: Summary of parameters for the SSDMD model used for each data set.

3.2 SSDMD Model Performance

We tested the SSDMD method on 30 randomly chosen 12-day periods in the BC840 and RO041 data sets. Each of these stations contained several large gaps in their data which were not used in the random start times as one cannot fit an SSDMD model without contiguous data. Even though standard DMD methods will work for arbitrary snapshots of data $(\mathbf{x}, \mathbf{y})$, where $\mathbf{y} = \mathbf{K}\mathbf{x}$, the wavelet decompositions used in SSDMD require a regular measurement cadence, i.e., the data snapshots are always δt time apart.

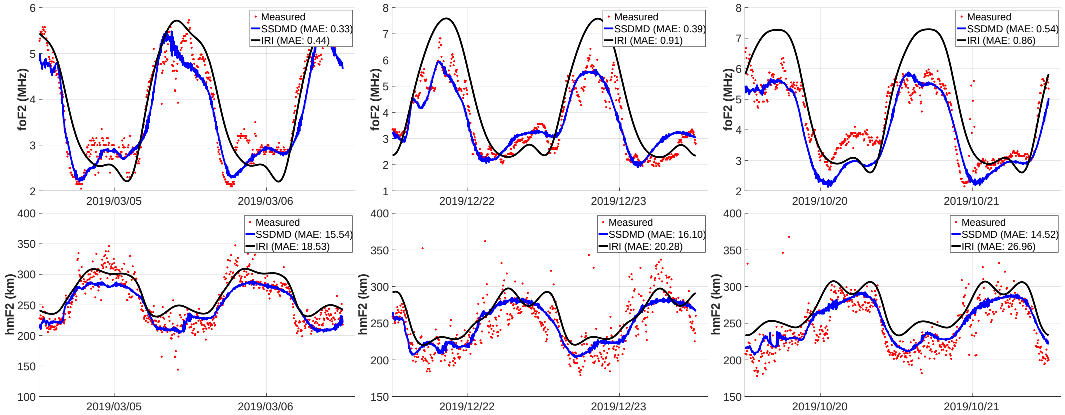


Figure 10: SSDMD forecasts of foF2 (top panels) and hmF2 (bottom panels) for the BC840 sounding station for randomly chosen starting times in 2019. The MAE is provided for both the SSDMD and IRI forecasts.

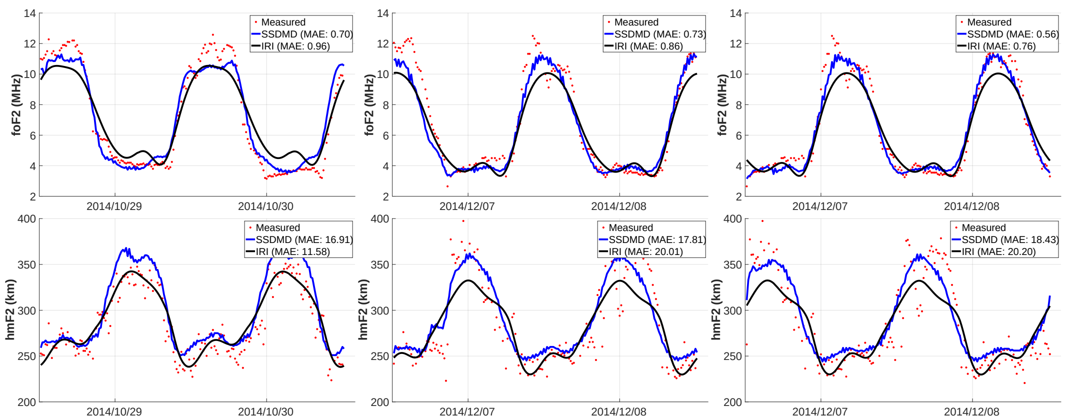


Figure 11: SSDMD forecasts of foF2 (top panels) and hmF2 (bottom panels) for the RO041 sounding station for randomly chosen starting times in 2014. The MAE is provided for both the SSDMD and IRI forecasts.

For each random 12-day period, the first 10 days were used for fitting an SSDMD model and the remaining 2 days for testing a 48-hour forecast of the foF2 and hmF2 parameters. Figures 10 and 11 show these test forecast periods for 3 of the 30 randomly chosen times in each of the BC840 and RO041 data sets. The SSDMD and IRI predictions for the F-layer parameters, along with the measured values from the sounder, are presented for each. From these, we see that SSDMD captures some smaller-scale fluctuations in the parameters that are commonly lost in climatological models due to extreme averaging over monthly and seasonal variations. The mean absolute error (MAE) is provided for each forecast. While, in general, the SSDMD MAE shows modest improvements over IRI for BC840 in 2019, it is not always the case, as we can see in the hmF2 forecast for RO041 in 2014. However, in the cases where SSDMD does perform worse than IRI, it is still relatively close considering how little data is used to generate the forecast.

Figures 12 and 13 provide scatter plots and histograms of the foF2 modeled vs. measured forecasts for BC840 and RO041, respectively. The histograms are given to illustrate the shapes of the total model error distributions. The area of each bin simply rep-

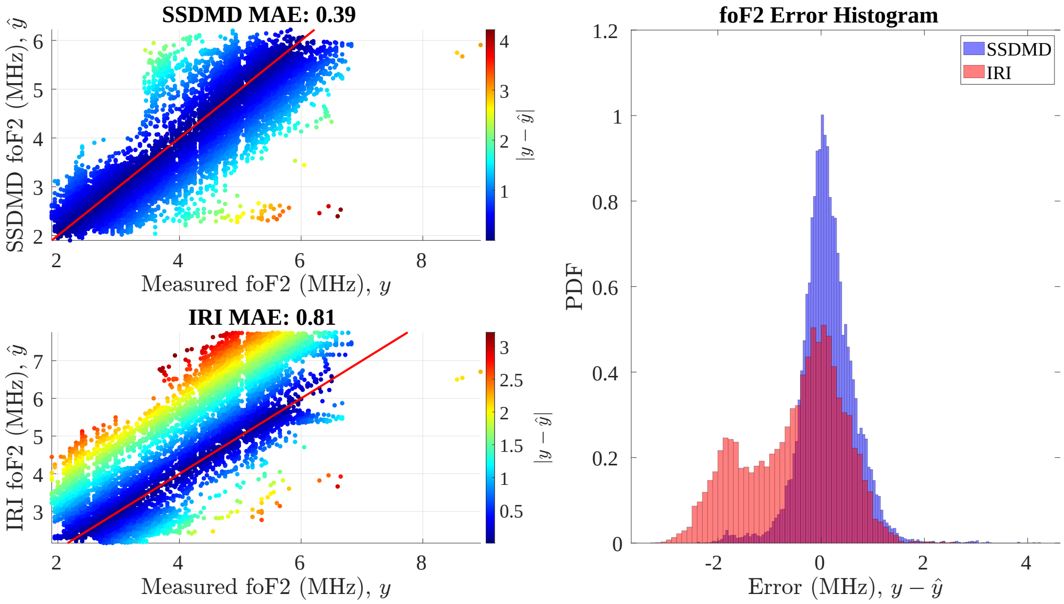


Figure 12: Forecasted vs. measured foF2 parameter scatter plots for the SSDMD (top left) and IRI (bottom left) models for the BC840 station in 2019. The total MAE for each model is given above their respective scatter plot. Histograms (right) provide estimates of the total model error distributions.

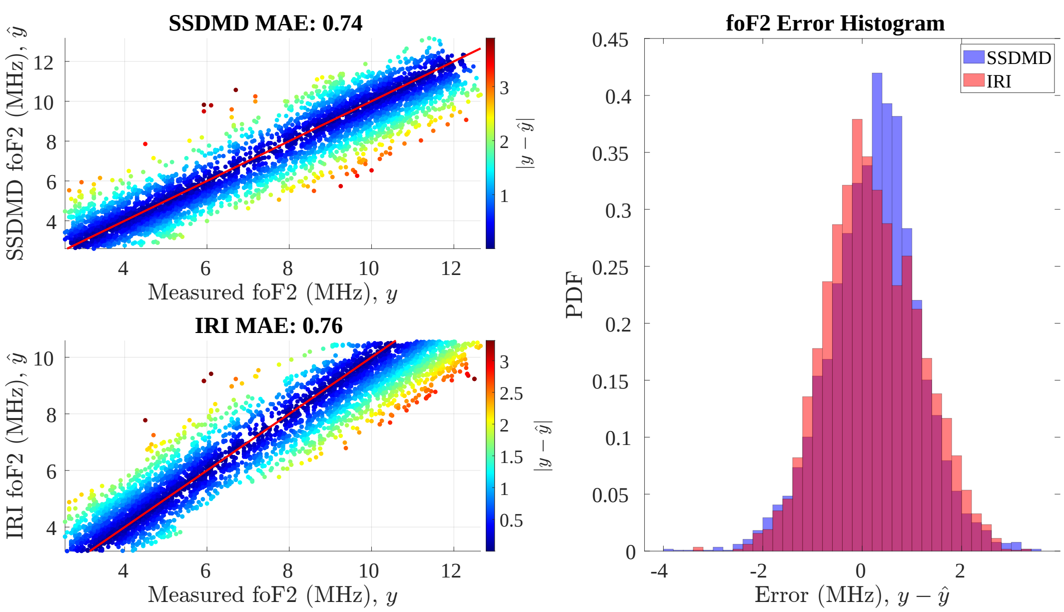


Figure 13: Forecasted vs. measured foF2 parameter scatter plots for the SSDMD (top left) and IRI (bottom left) models for the RO041 station in 2014. The total MAE for each model is given above their respective scatter plot. Histograms (right) provide estimates of the total model error distributions.

resents the relative number of model errors within that interval over all 48-hour forecast test periods. Note that SSDMD forecasts perform markedly better on the BC840 data set, with IRI producing a significant bimodal error distribution. This may point toward limitations in SSDMD's applicability during periods of high solar activity. Figure 8 shows a significant seasonal variation in the foF2 parameter of the RO041 station. Applying SSDMD to longer time series to capture seasonal and solar cycle trends will be a topic of future study. Furthermore, in the context of short-term forecasts, SSDMD's reliance on the fit of the $\mathbf{K}_o$ matrix to advance any data point one time-step into the future benefits from higher measurement cadences. In addition, as the time resolution of sounder measurements increases, a wider spectrum of geophysical noise will be observed, and thus, SSDMD's ability to identify couplings between dominant scales becomes more pronounced.

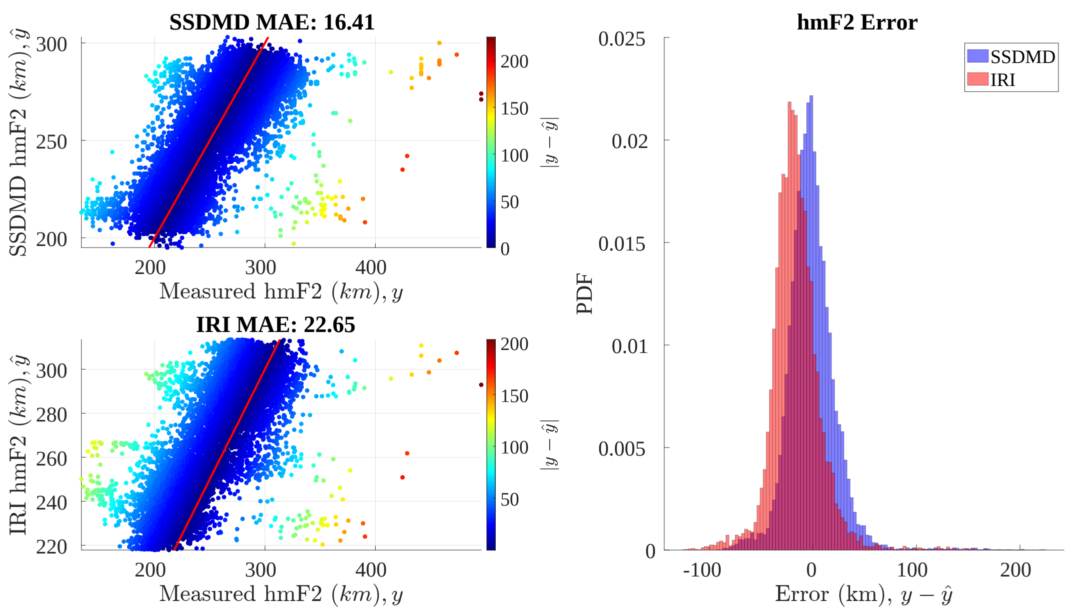


Figure 14: Forecasted vs. measured hmF2 parameter scatter plots for the SSDMD (top left) and IRI (bottom left) models for the BC840 station in 2019. The total MAE for each model is given above their respective scatter plot. Histograms (right) provide estimates of the total model error distributions.

Figures 14 and 15 give similar scatter plots and histograms for the hmF2 parameter for the BC840 and RO041 stations, respectively. With hmF2, we find the model error distributions for both SSDMD and standard IRI to be very similar. However, SSDMD provides a slight bias correction over IRI for the BC840 data set. While the hmF2 MAE for SSDMD on the RO041 data is worse than IRI, its performance is still quite close, given the relatively small amount of data used to generate the forecast.

The SSDMD model was run on the GA762 station data set to illustrate its use on data streams from higher latitudes. GA762 is at a latitude of 62.38°N and is the site of the High-frequency Active Auroral Research Program (HAARP) (Bailey & Worthington, 2000), a valuable ionospheric-thermospheric research instrument used in a variety of fundamental and experimental physics applications (Bell, 2001; Bernhardt et al., 2009). Improved forecasts of the foF2 and hmF2 parameters continue to play a critical role in high-frequency radio experimentation and modeling. The use of a lightweight and adaptive forecast like SSDMD for real-time operations may be explored in future work, but in this paper we use this station to provide validation of our method in these high-latitude regions. Figures 16 and 17 give forecasts of foF2 and hmF2 and visualizations of the full EDP reconstructions for this station.

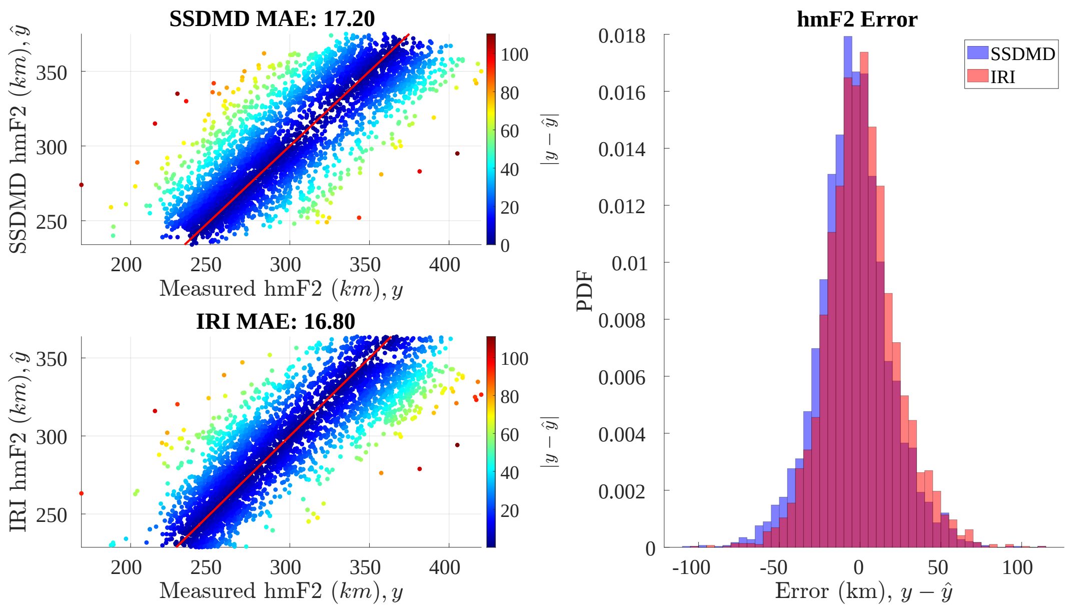


Figure 15: Forecasted vs. measured hmF2 parameter scatter plots for the SSDMD (top left) and IRI (bottom left) models for the RO041 station in 2014. The total MAE for each model is given above their respective scatter plot. Histograms (right) provide estimates of the total model error distributions.

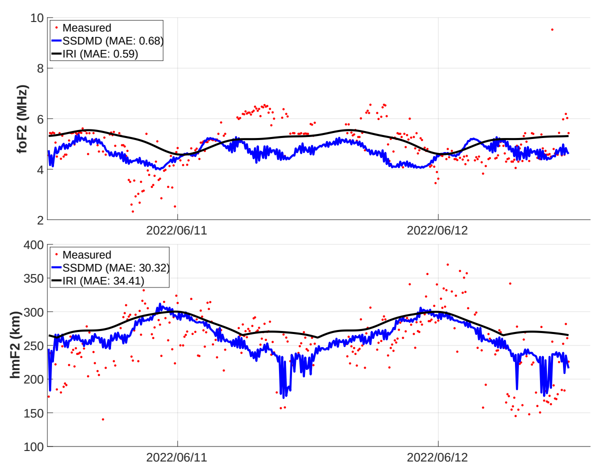


Figure 16: SSDMD 2-day forecast of the foF2 (top) and hmF2 (bottom) parameters for the GA762 station with IRI predictions. MAE values for both models are provided in the legend.

Lastly, Figures 18 and 19 demonstrate the SSDMD model in a low-latitude environment. Figure 19 illustrates the dramatic oscillations of the hmF2 as compared with the mid- and high-latitude stations. The presence of complex physical processes like the equatorial plasma fountain (MacDougall, 1969; Balan et al., 2018) induce categorically more complex dynamics in the EDP time series than observed at mid-latitudes. Still, we find SSDMD can fit a model that improves the MAE for both foF2 and hmF2 compared to IRI.

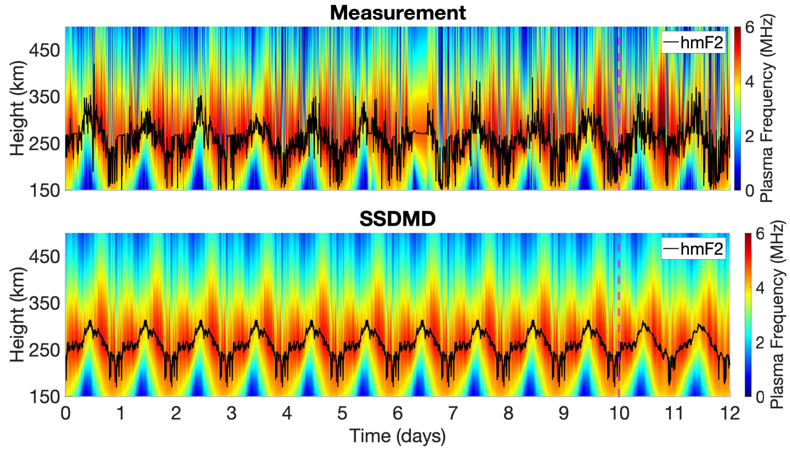


Figure 17: SSDMD full EDP time series reconstruction and 2-day forecast of the GA762 station during a 15-day period in 2022. The vertical dotted magenta line indicates the transition from fitting data to test data, and the solid black line follows the hmF2 parameter computed using the EDP time series.

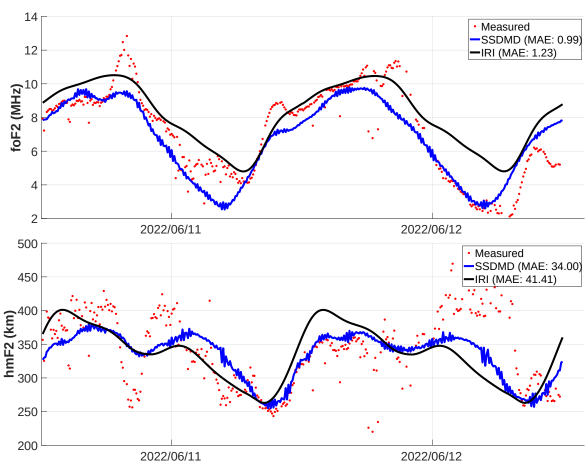


Figure 18: SSDMD 2-day forecast of the foF2 (top) and hmF2 (bottom) parameters for the GU513 station with IRI predictions. MAE values for both models are provided in the legend.

In addition to the MAE statistics presented for each station, Tables 3 and 4 give summaries of root-mean-squared error (RMSE) and mean absolute percentage error (MAPE) for all foF2 and hmF2 forecasts, respectively. We find that SSDMD either outperforms or closely matches a standard IRI forecast for both foF2 and hmF2 for the data sets presented. While significant improvement in the IRI forecast can be made by tweaking coefficients within the model or even through the assimilation of real time data, SSDMD provides an easily implementable fitting method that can adapt to new data in real-time. Moreover, adjusting the parameters within IRI will not always improve its forecast accuracy, as one does not know in which direction to adjust parameters until observations of the ionosphere are made.

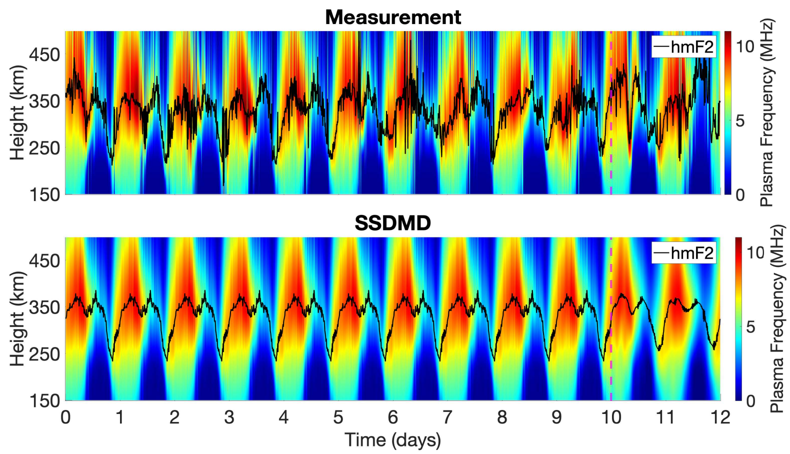


Figure 19: SSDMD full EDP time series reconstruction and 2-day forecast of the GU513 station during a 15-day period in 2022. The vertical dotted magenta line indicates the transition from fitting data to test data and the solid black line follows the hmF2 parameter computed using the EDP time series.

foF2 Forecast Errors						
	RMSE		MAE		MAPE	
Station	SSDMD	IRI	SSDMD	IRI	SSDMD	IRI
BC840	**0.54**	1.06	**0.39**	0.81	**10.08**	21.54
RO041	**0.93**	0.95	**0.74**	0.76	11.54	**11.44**
GA762	0.91	**0.81**	0.68	**0.59**	13.59	**13.18**
GU513	**1.26**	1.57	**0.99**	1.23	**16.02**	26.30

Table 3: Summary of foF2 error statistics for all stations using SSDMD and IRI.

4 Conclusions and Future Directions

We presented the standard DMD algorithm and formalized extensions that account for oscillations at multiple scales within measured data. Wavelet decompositions along each spatial dimension separated various scales within the time series that may otherwise appear as noise and will often preclude a standard DMD approach. For each of the scales, an affiliated reconstruction of the EDP time series was generated. Subsequent correlation analysis across the time scales then showed how we may recombine specific scales to preserve strong dynamic couplings between them in their one-step correlation. We called these correlated scales the connected components of the model. We performed an averaging step for each connected component by computing the mean over 24-hour time lags. This process denoises the data without erroneously removing oscillations from the original EDP signal that may initially appear as noise. Computing DMD on the connected

hmF2 Forecast Errors						
	RMSE		MAE		MAPE	
Station	SSDMD	IRI	SSDMD	IRI	SSDMD	IRI
BC840	**23.03**	28.68	**16.41**	22.65	**6.72**	9.72
RO041	22.72	**22.15**	17.20	**16.80**	5.91	**5.66**
GA762	**38.30**	45.87	**30.32**	34.41	**13.08**	16.02
GU513	**45.43**	54.74	**34.00**	41.41	**9.84**	11.90

Table 4: Summary of hmF2 error statistics for all stations using SSDMD and IRI.

components individually alleviates the problem of having large single-step gradients in the measurement data that would prevent DMD from fitting any stable modes. With each connected component, we produced a set of DMD eigenvalues and modes that summed coherently to form the SSDMD model. The final foF2 and hmF2 forecasts were then determined from the predicted EDPs.

SSDMD is one among many recent attempts to improve short-term forecasts of the foF2 and hmF2 parameters (cf. Perrone & Mikhailov, 2022; Wang et al., 2020; Tsagouri et al., 2018; Mikhailov & Perrone, 2014; Zhang et al., 2014). While other methods generally treat past foF2 or hmF2 measurements as inputs to the model, SSDMD instead uses the full EDP. The number of DMD modes is limited by the initial dimensionality of the data; see Equation 8. Therefore, if the data used to generate the model only consisted of the foF2 and hmF2 parameters, we would be restricted to a maximum of two eigenvalues. Instead, using the high-dimensional EDP from the sounder gives our method far richer spectral properties.

We note that all profile data in this study are autoscaled. This is an inherent data limitation as there are no widely available manually scaled data sets that are of a size suitable for statistical analysis. However, future studies with SSDMD and manually scaled data may reveal additional insights into the spatial and temporal distributions of fluctuations. Despite using autoscaled EDPs to construct the SSDMD models, our forecast errors reflect the method's predictions of foF2 and hmF2 and not the full profile.

The SSDMD algorithm is computationally efficient compared to physics-based models such as TIME-GCM or SAMI3, fitting a model and simulating a 5-minute resolution, 2-day forecast on the order of seconds using a single core on a consumer laptop. Therefore, SSDMD is lightweight enough to be updated in near-real-time as additional data are obtained, and it adapts to different measurement cadences without any changes to the model parameters. Additionally, SSDMD requires far less data to generate and update than empirical models like IRI or assimilation models like IRI-Real-Time-Assimilative-Mapping (IRTAM) (Galkin et al., 2012) and the Global Assimilation of Ionospheric Measurements (GAIM) model (Schunk et al., 2004). With limited observations, as is the case with a single vertical ionosonde, SSDMD can produce reasonable forecasts of the average profile dynamics in the low, mid, and high latitudes. With high enough measurement cadence, the method should produce reliable short-term forecasts during periods of either solar maximum or solar minimum. A final added benefit of the SSDMD approach is the model has only four major hyperparameters, see Table 2, making it relatively simple to tune when necessary.

SSDMD fits a linear model to an expansion of full EDP time series and thus may be seen as an autoregressive approach to forecasting foF2 and hmF2, and the simplicity of the approach makes it accessible to a wide range of operational and research applications. Still, the method is not without its limitations, as SSDMD does not account for any external driving forces such as solar activity, tidal forcing, or geomagnetic activity. As such, model forecast accuracy is highly dependent on there being strong correlations between the measurement and forecast periods at each time of day. Predicting anomalous events in the data is not possible without the inclusion of driving forces. Extending SSDMD further to incorporate external forcing is the topic of future development and, combined with longer measurement series, could allow for a significant increase in forecast accuracy over much longer prediction windows. The DMD method can be modified to include control variables (Proctor et al., 2016), and in Mehta et al. (2018) a version of this method was implemented for a global model to great effect. Nevertheless, this model was fit using simulated data, whereas SSDMD aims to address the multiscale nature of measured EDPs. For this reason, the applicability of SSDMD to periods of prolonged or recurrent F-layer perturbations during quiet geomagnetic conditions may also be explored in future work. These disturbances can induce long-lived deviations in foF2 and hmF2 with magnitudes that far exceed climatology (Perrone et al., 2020; Zawdie et al., 2020) which would not necessarily be captured by empirical models with drivers derived from geomagnetic and solar indices.

While the method was developed for one-dimensional observations of the ionosphere at a single sounder station, in future work, data from the global network of sounders may be used. However, a global model will require fitting additional spatial expansion functions to interpolate between the stations. Finally, data spanning longer time periods may also be used to extract seasonal and solar cycle dynamics. The method of SSDMD is ultimately not limited to ionospheric prediction, and it should be adaptable not only to other space weather domains, but many other systems that involve low-dimensional dynamics embedded in high-dimensional, multiscale observations.

5 Data Availability Statement

The code used in this study is openly available at `https://github.com/JayLago/SSDMD-Ionosphere` or at the permanent release link, `https://doi.org/10.5281/zenodo.7109436`. The data used was obtained through the LGDC, `https://giro.uml.edu/didbase/`, using the SAO Explorer program, for which we are grateful to the developers and maintainers. The authors would like to thank Dr. Terrance Bullet from the National Centers for Environmental Information, NOAA, as well as Dr. Ivan Galkin from the LGDC for the data from the Boulder, CO Digisonde station they continue to collect and make available.

Appendix A Koopman Mode Analysis

Dynamic Mode Decomposition may be seen as a finite-dimensional approximation to the Koopman operator (Koopman, 1931). The Koopman operator demonstrates how the equations for a generic nonlinear dynamical system may be rewritten as a linear infinite-dimensional operator acting on measurement functions of the system. This begins by considering a generic dynamical system,

$$\frac{d}{dt}\mathbf{y}(t) = f(\mathbf{y}(t)), \quad \mathbf{y}(0) = \mathbf{y}_0 \in \mathcal{M} \subseteq \mathbb{R}^{N_s}, \tag{A1}$$

where $\mathcal{M}$ is some connected, compact subset of $\mathbb{R}^{N_s}$ and define an *observable*, $g(\mathbf{y}(t))$, such that $g : \mathcal{M} \mapsto \mathbb{C}$. Denoting the affiliated flow, $\mathbf{y}(t) = S(t; \mathbf{y})$, we may rewrite the system using the *Koopman operator*, $\mathcal{K}^t$,

$$\mathcal{K}^t g(\mathbf{y}) = g\left(S(t; \mathbf{y})\right). \tag{A2}$$

We see $\mathcal{K}^t$ is linear since

$$\begin{aligned} \mathcal{K}^t(\alpha g_1(\mathbf{y}) + \beta g_2(\mathbf{y})) &= \alpha g_1(S(t; \mathbf{y})) + \beta g_2(S(t; \mathbf{y})) \\ &= \alpha \mathcal{K}^t g_1(\mathbf{y}) + \beta \mathcal{K}^t g_2(\mathbf{y}). \end{aligned} \tag{A3}$$

Following (Alford-Lago et al., 2022), we see that with some basic assumptions, i.e. if we choose observables such that they are square-integrable and suppose $\mathcal{M}$ is invariant with respect to the flow, we have simplified a problem of determining some unknown nonlinear function $f(\mathbf{y}(t))$ to one of finding an eigendecomposition of the linear operator, $\mathcal{K}^t$. Moreover, by finding the Koopman eigenfunctions

$$\{\phi_j\}_{j=1}^{\infty} \tag{A4}$$

and affiliated eigenvalues

$$\{\lambda_j\}_{j=1}^{\infty}, \tag{A5}$$

where

$$\mathcal{K}^t \phi_j = e^{t\lambda_j}\phi_j, \quad j \in \{1, 2, \ldots\}, \tag{A6}$$

then we have a modal decomposition for any other observable, g, so that

$$g(\mathbf{y}) = \sum_{j=1}^{\infty} c_j \phi_j(\mathbf{y}), \tag{A7}$$

and we can track the evolution of $g(\mathbf{y})$ along the flow with the formula,

$$\mathcal{K}^t g(\mathbf{y}) = \sum_{j=1}^{\infty} c_j e^{t\lambda_j} \phi_j(\mathbf{y}). \tag{A8}$$

See (Budisić et al., 2012) and (Mezić, 2019) for more in-depth treatments of the Koopman operator and its properties, (Mezić, 2005; Kutz et al., 2016) for deeper connections between DMD and Koopman, and (Schmid, 2010; Tu et al., 2014; Williams et al., 2015) for additional details on the DMD algorithm and its variations. We point out that the Koopman operator is most naturally formulated with respect to Lagrangian data while in this work we focus on analyzing Eularian data, that is to say, we assume the $\mathbf{y}_j$ observations in our data stream are measurements of the EDP at fixed positions in altitude. Were one to develop effective Euler-to-Lagrangian maps for the data sets studied herein, this would open up a wider range of tools related to the DMD method. This is a subject for future research.

Appendix B Pseudocode Algorithm

The complete SSDMD method is summarized in Algorithm 1. We assume familiarity with standard numerical methods for computing the reduced Singular Value Decomposition (SVD), eigenvalue decomposition, solving an initial value problem, and computing 1-dimensional wavelet decompositions. When computing the mean profiles over 24-cycles, use Equation 24. The algorithm returns the reconstructed time series of the input data along with the DMD eigenvalues, modes, and eigenfunctions.

Acknowledgments

This work was supported by the Naval Information Warfare Center Pacific (NIWC Pacific) and the Office of Naval Research (ONR).

Algorithm 1: SSDMD

Data: $\mathbf{Y} \in \mathbb{R}^{N_S \times N_T}$ such that each column, $\mathbf{y}_i \in \mathbb{R}^{N_S}$, is an observation of the system δt time from $\mathbf{y}_{i-1}$.

Result: $\hat{\mathbf{Y}}, \mathbf{W}, \mathbf{\Lambda}, \mathbf{\Phi}$

Initialize: set DMD threshold $c_{\mathrm{dmd}} > 0$, and correlation threshold $c_{\mathrm{corr}} > 0$.

begin

- $\tilde{\mathbf{Y}} \longleftarrow discreteWaveletDecomposition(\mathbf{Y})$
- $\tilde{\mathbf{Y}}^{\mathrm{C}}, N_{\mathrm{C}} \longleftarrow correlatedConnectedComponents(\tilde{\mathbf{Y}}, c_{tr})$
- **for** n=1 $\dots N_C$ **do**
  - $\bar{\mathbf{Y}}_n^{\mathrm{C}} \longleftarrow meanDailyCycles(\tilde{\mathbf{Y}}_n^{\mathrm{C}})$
  - $\bar{\mathbf{Y}}_{n,-}^{\mathrm{C}} \longleftarrow \left[\bar{\mathbf{y}}_{n,1}^{\mathrm{C}}\ \bar{\mathbf{y}}_{n,2}^{\mathrm{C}}\ \cdots \bar{\mathbf{y}}_{n,m-1}^{\mathrm{C}}\right]$
  - $\bar{\mathbf{Y}}_{n,+}^{\mathrm{C}} \longleftarrow \left[\bar{\mathbf{y}}_{n,2}^{\mathrm{C}}\ \bar{\mathbf{y}}_{n,3}^{\mathrm{C}}\ \cdots \bar{\mathbf{y}}_{n,m}^{\mathrm{C}}\right]$
  - $\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}^{\dagger} \longleftarrow reducedSVD(\bar{\mathbf{Y}}_{n,-}^{\mathrm{C}}, C_{dmd})$
  - $\mathbf{K} \longleftarrow \bar{\mathbf{Y}}_{n,+}^{\mathrm{C}} \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{U}^{\dagger}$
  - $\mathbf{W}_n, \mathbf{\Lambda}_n \longleftarrow eigenvalueDecomposition(\mathbf{K})$
  - $\mathbf{\Phi}_n \longleftarrow solveIVP(\mathbf{W}_n, \bar{\mathbf{Y}}_{n,-}^{\mathrm{C}})$
  - $\hat{\mathbf{Y}}_n \longleftarrow \mathbf{W}_n \mathbf{\Lambda}_n \mathbf{\Phi}_n$
- $\hat{\mathbf{Y}} \longleftarrow \sum_{n=1}^{N_{\mathrm{C}}} \hat{\mathbf{Y}}_n$
- $\mathbf{W} \longleftarrow [\mathbf{W}_1\ \mathbf{W}_2\ \cdots\ \mathbf{W}_n]$
- $\mathbf{\Lambda} \longleftarrow [\mathbf{\Lambda}_1\ \mathbf{\Lambda}_2\ \cdots\ \mathbf{\Lambda}_n]$
- $\mathbf{\Phi} \longleftarrow [\mathbf{\Phi}_1\ \mathbf{\Phi}_2\ \cdots\ \mathbf{\Phi}_n]$

The authors would like to especially thank Dr. Douglas Drob from the Naval Research Laboratory Space Science Division for his insights and perspective on the use of this method.

References

Alford-Lago, D. J., Curtis, C. W., Ihler, A. T., & Issan, O. (2022). Deep learning enhanced dynamic mode decomposition. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *32*(3), 033116. Retrieved from `https://doi.org/10.1063/5.0073893` doi: 10.1063/5.0073893

Araujo-Pradere, E. A., Fuller-Rowell, T. J., Codrescu, M. V., & Bilitza, D. (2005). Characteristics of the ionospheric variability as a function of season, latitude, local time, and geomagnetic activity. *Radio Science*, *40*(RS5009).

Bailey, P. G., & Worthington, N. C. (2000). History and applications of haarp technologies: The high frequesncy active auroral research program. In (p. 1317-1322). Proceedings of the Intersociety Energy Conversion Engineering Conference.

Balan, N., Liu, L., & Le, H. (2018). A brief review of equatorial ionization anomaly and ionospheric irregularities. *Earth and Planetary Physics*, *2*, 257-275.

Bell, T. F. (2001). *Characterization of the auroral electrojet and the ambient and modified d region for haarp using long-path vlf diagnostics* (Report No. AFR-LVSTR20011573 // ADA405592). Stanford University, Space Telecommunications and Radioscience Lab.

Bernhardt, P. A., Selcher, C. A., Lehmberg, R. H., Rodriguez, S., McCarrick, J., & Frazer, G. (2009). Determination of the electron temperature in the modified ionosphere over haarp using the hf pumped stimulated brillouin scatter (sbs) emission lines. *Annales Geophysicae*, *27*, 4409-4427.

Bibl, K., Reinisch, B. W., & Kitrosser, D. F. (1981, Dec). *General description of the compact digital ionospheric sounder digisonde 256. revised.* (Tech. Rep. No. ADA277181). Defense Technical Information Center.

Bilitza, D. (2001, March/April). International reference ionosphere 2000. *Radio Science*, *36*(2), 261-275.

Bilitza, D. (2018). Iri the international standard for the ionosphere. *Advances in Radio Science*, *16*, 1-11.

Booker, H. G. (1979). The role of acoustic gravity waves in the generation of spread-f and ionospheric scintillation. *J. Atmos. Terr. Phys.*, *41*, 501-515.

Budden, K. G. (1985). *The propagation of radio waves: The theory of radio waves of low power in the ionosphere and magnetosphere.* Cambridge: Cambridge University Press.

Budisić, M., Mohr, R., & Mezić, I. (2012). Applied koopmanism. *Chaos*, *22*(047510).

Curtis, C., Carretero-Gonzalez, R., & Polimeno, M. (2019). Characterizing coherent structures in Bose-Einstein condensates through dynamic-mode decomposition. *Phys. Rev. E*, *99*, 062215.

Davies, K. (1990). *Ionospheric radio.* London: Peregrinus.

Dickinson, R. E., Ridley, E. C., & Roble, R. G. (1981). A three-dimensional general circulation model of the thermosphere. *J. Geophys. Res.*, *86*, 1499-1512.

Dylewsky, D., Tao, M., & Kutz, J. N. (2019). Dynamic mode decomposition for multiscale nonlinear physics. *Phys. Rev. E*, *99*(6), 063311–063311.

Fagre, M., Zossi, B. S., Chum, J., Yigit, E., & Elias, A. G. (2019). Ionospheric high frequency wave propagation using different iri hmf2 and fof2 models. *Journal of Atmospheric and Solar-Terrestrial Physics*, *196*, 105141.

Galkin, I. A., & Reinisch, B. W. (2008, 01). The new artist 5 for all digisondes. *Ionosonde Network Advis. Group Bull.*, *69*.

Galkin, I. A., Reinisch, B. W., Huang, X., & Bilitza, D. (2012). Assimilation of giro data into a real-time iri. *Radio Science*, *47*(RS0L07).

Grzesiak, M., Cesaroni, C., Spogli, L., Franceschi, G. D., & Romano, V. (2018). Regional short-term forecasting of ionospheric tec and scintillation. *Radio Science*, *53*, 1254-1268.

Huba, J. D., Joyce, G., & Fedder, J. A. (2000). Sami2 is another model of the ionosphere (sami2): a new low-latitude ionosphere model. *J. Geophys. Res. Space Phys.*, *105*(A10), 23035-23053.
Huba, J. D., & Krall, J. (2013). Modeling the plasmasphere with sami3. *Geophysical Research Letters*, *40*, 6-10.
Killeen, T. (1987). Energetics and dynamics of the earth's thermosphere. *Rev. Geophys.*, *25*(3), 433-454.
Koopman, B. (1931). Hamiltonian systems and transformation in Hilbert space. *Proc. Nat. Acad. Sci.*, *17*, 315–318.
Kutz, J. N., Brunton, S. L., Brunton, B. W., & Proctor, J. L. (2016). *Dynamic mode decomposition: Data-driven modeling of complex systems.* Philadelphia, PA: SIAM.
Kutz, J. N., Fu, X., & Brunton, S. L. (2015). Multi-resolution dynamic mode decomposition. *arXiv*.
Li, Q., Liu, L., He, M., Huang, H., Zhong, J., Yang, N., ... Cui, J. (2021, April). A global empirical model of electron density profile in the f region ionosphere basing on cosmic measurements. *Space Weather*, *19*(4).
Liu, C., Zhang, M.-L., Wan, W., Liu, L., & Ning, B. (2008). Modeling m(3000)f2 based on empirical orthogonal function analysis method. *Radio Science*, *43*(43).
Liu, H. L. (2016). Variability and predictability of the space environment as related to lower atmosphere forcing. *Space Weather*, *14*, 634-658.
MacDougall, J. W. (1969). The equatorial ionospheric anomaly and the equatorial electrojet. *Radio Science*, *4*(9), 805-810.
Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7).
Mallat, S. (2009). *A wavelet tour of signal processing.* New York, NY: Academic Press.
Mehta, P. M., & Linares, R. (2017). A methodology for reduced order modeling and calibration of the upper atmosphere. *Space Weather*, *15*, 1270-1287.
Mehta, P. M., Linares, R. L., & Sutton, E. K. (2018). A quasi-physical dynamic reduced order model for thermospheric mass density via hermitian space-dynamic mode decomposition. *Space Weather*, *16*, 569-588.
Mezić, I. (2005). Spectral properties of dynamical systems, model reduction, and decompositions. *Nonlinear Dyn.*, *41*, 309–325.
Mezić, I. (2019). Spectrum of the koopman operator, spectral expansions in functional spaces, and state-space geometry. *Journal of Nonlinear Science*, *30.5*, 2091–145.
Mikhailov, A. V., & Perrone, L. (2014). A method for fof2 short-term (1–24 h) forecast using both historical and real-time fof2 observations over european stations: Euromap model. *Radio Science*, *49*, 253-270.
Perrone, L., & Mikhailov, A. V. (2022). Srf2—a short-term (1–24)h fof2 prediction method. *Space Weather*, *20*.
Perrone, L., Mikhailov, A. V., & Nusinov, A. A. (2020). Daytime mid-latitude f2-layer q-disturbances: A formation mechanism. *Scientific Reports*, *10*(1), 9997.
Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. (2002). Nrlmsise-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *J. Geophys. Res.*, *107*(A12), 1468.
Proctor, J. L., Brunton, S. L., & Kutz, J. N. (2016). Dynamic mode decomposition with control. *SIAM J. Appl. Dyn. Sys.*, *15*(1), 142-161.
Ratcliffe, J. A. (1959). *The magnetoionic theory and its applications to the ionosphere. a monograph.* London: Cambridge University Press.
Reinisch, B. W., & Galkin, I. A. (2011). Global ionospheric radio observatory (giro). *Earth, Planets, and Space*, *63*(4), 377-81.

Richmond, A. D., Ridley, E. C., & Roble, R. G. (1992). A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophys. Res. Lett.*, *19*, 601-604.
Roble, R. G. (1995). Energetics of the mesosphere and thermosphere. *AGU Geophysical Monograph*, *87*, 1-22.
Roble, R. G., & Ridley, E. C. (1994). A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (time-gcm): equinox solar cycle minimum simulations (30-500 km). *Geophys. Res. Lett.*, *21*, 417-420.
Schmid, P. (2010). Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.*, *656*, 5–28.
Schunk, R. W., Scherliess, L., Sojka, J. J., Thompson, D. C., Anderson, D. N., Codrescu, M., . . . Howe, B. M. (2004). Global assimilation of ionospheric measurements (gaim). *Radio Science*, *39*.
Stanislawska, I., & Zbyszynski, Z. (2001). Forecasting of the ionospheric quiet and disturbed fof2 values at a single location. *Radio Science*, *36*(5), 1065-1071.
Tsagouri, I., Goncharenko, L., Shim, J. S., Belehaki, A., Buresova, D., & Kuznetsova, M. M. (2018). Assessment of current capabilities in modeling the ionospheric climatology for space weather applications: fof2 and hmf2. *Space Weather*, *16*, 1930-1945.
Tu, J., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., & Kutz, J. N. (2014). On dynamic mode decomposition: On dynamic mode decomposition: Theory and applications. *J. Comp. Dyn.*, *1*(2), 391-421.
Wang, J., Feng, F., & Ma, J. (2020). An adaptive forecasting method for ionospheric critical frequency of f2 layer. *Radio Science*, *55*.
Williams, M., Kevrekidis, I. G., & Rowley, C. W. (2015). A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. *J. Nonlin. Sci.*, *25*, 1307–1346.
Zawdie, K., Dhadly, M. S., McDonald, S. E., Sassi, F., Coker, C., & Drob, D. P. (2020). Day-to-day variability of the bottomside ionosphere. *J. Atmos. Terr. Phys.*, *205*, 105299.
Zhang, M.-L., Liu, C., Wan, W., Liu, L., & Ning, B. (2009). A global model of the ionospheric f2 peak height based on eof analysis. *Ann. Geophys.*, *27*, 3203-3212.
Zhang, M.-L., Liu, L., Wan, W., & Ning, B. (2014). An update global model of hmf2 from values estimated from ionosonde and cosmic/formosat-3 radio occultation. *Advances in Space Research*, *53*, 395-402.

Figure 6.

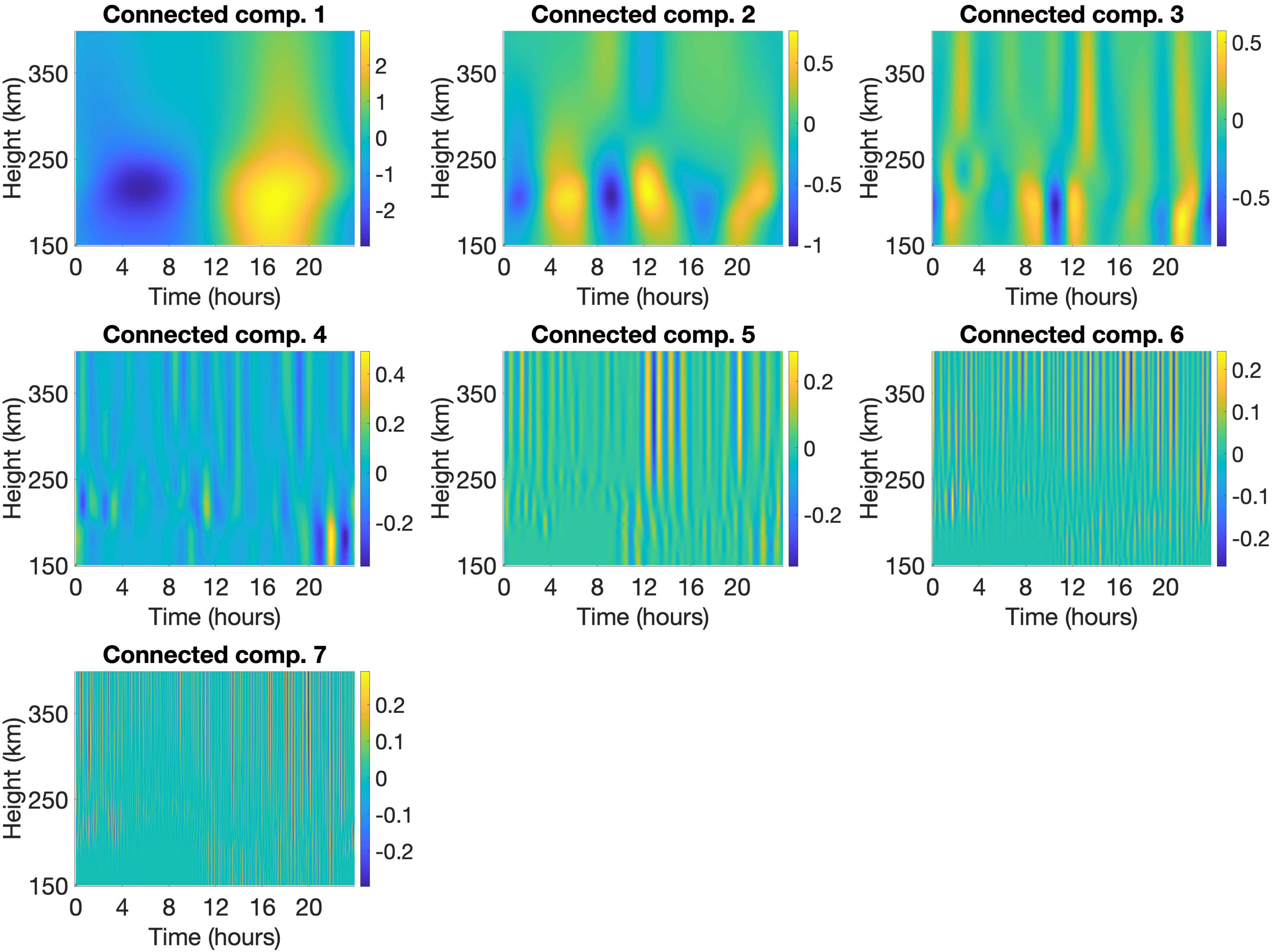

Connected comp. 1
Connected comp. 2
Connected comp. 3
Connected comp. 4
Connected comp. 5
Connected comp. 6
Connected comp. 7
Height (km)
Time (hours)

Figure 5.

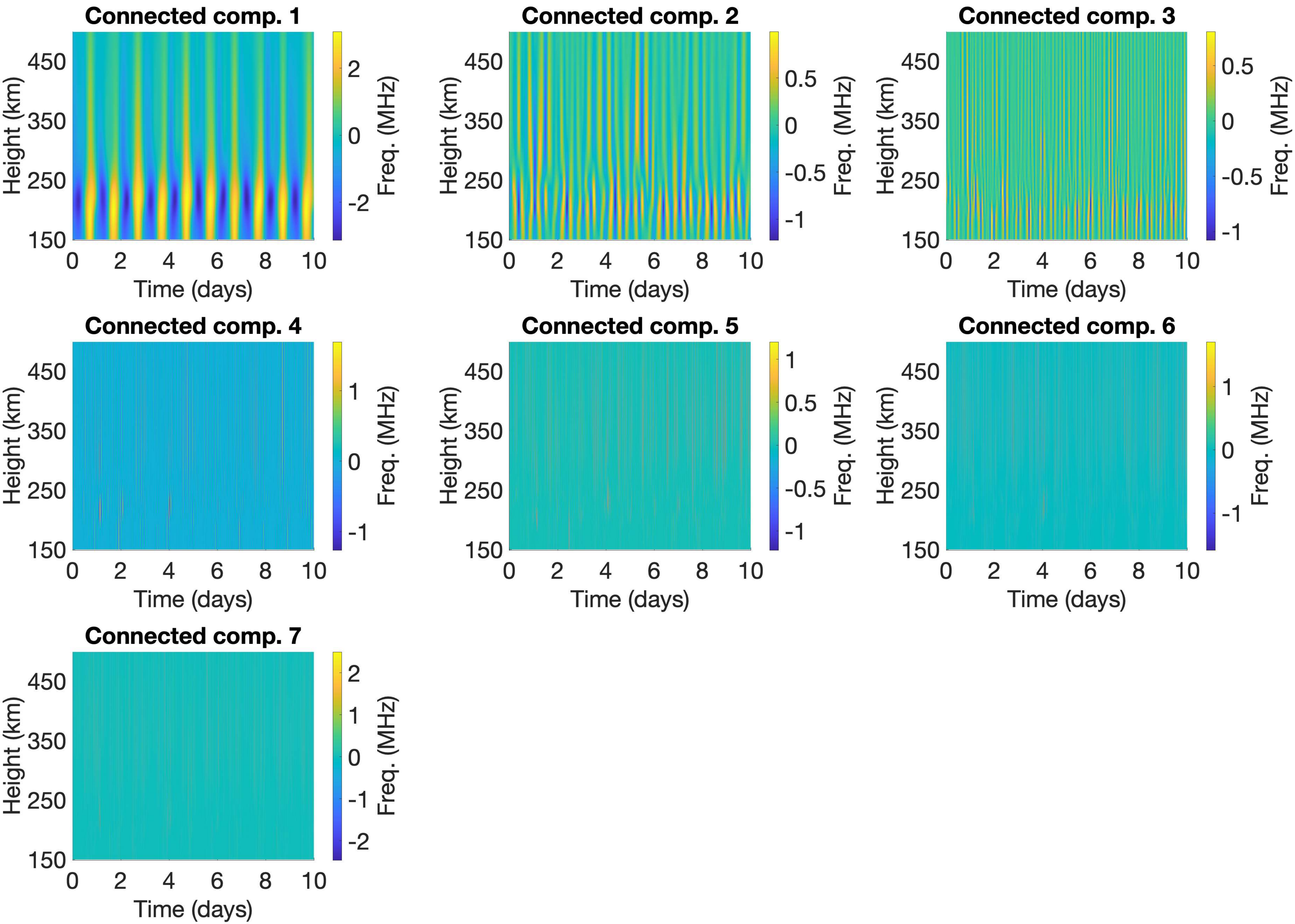

Connected comp. 1
Connected comp. 2
Connected comp. 3
Connected comp. 4
Connected comp. 5
Connected comp. 6
Connected comp. 7
Height (km)
Time (days)
Freq. (MHz)

Figure 4.

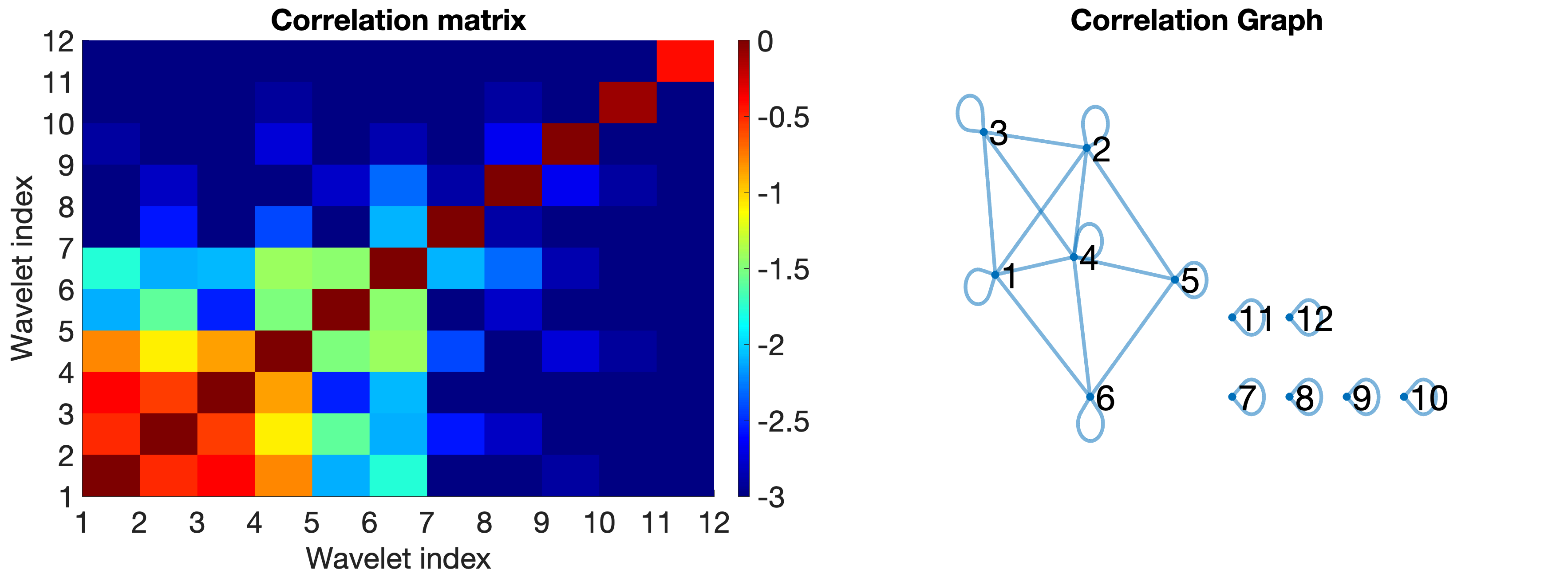
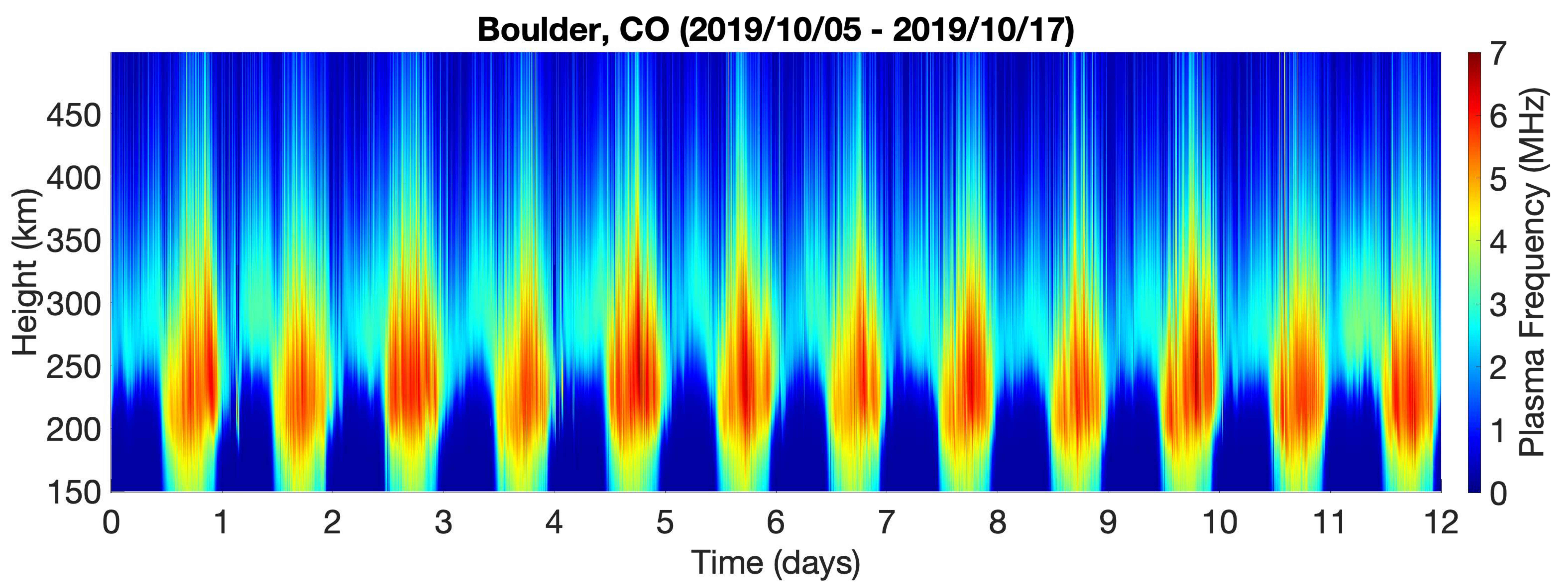

Correlation matrix
Wavelet index
Wavelet index
0
-0.5
-1
-1.5
-2
-2.5
-3
Correlation Graph
1
2
3
4
5
6
7
8
9
10
11
12

Figure 1.



Boulder, CO (2019/10/05 - 2019/10/17)
Height (km)
450
400
350
300
250
200
150
0
1
2
3
4
5
6
7
8
9
10
11
12
Time (days)
Plasma Frequency (MHz)
7
6
5
4
3
2
1
0

Figure 2.

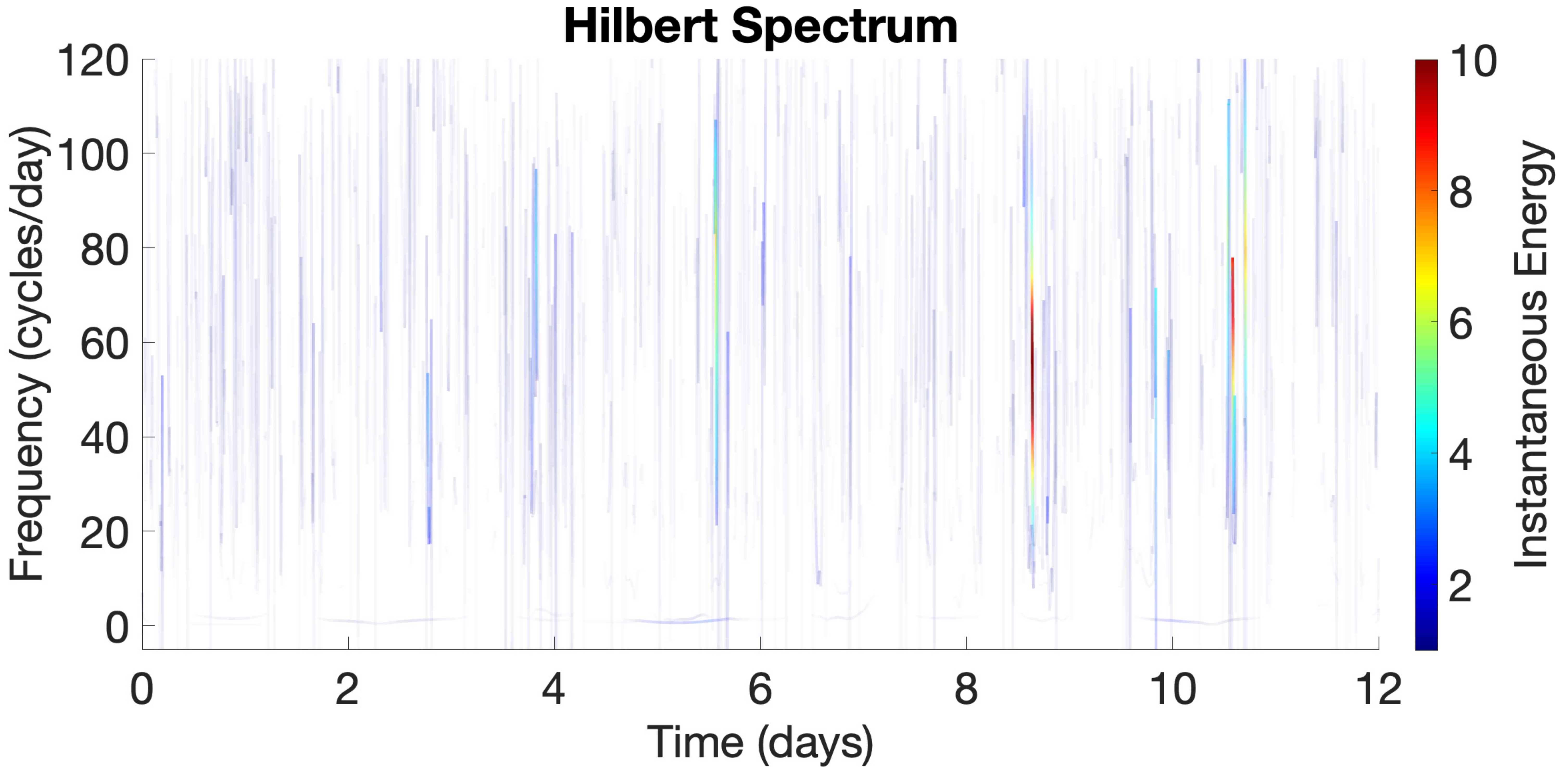
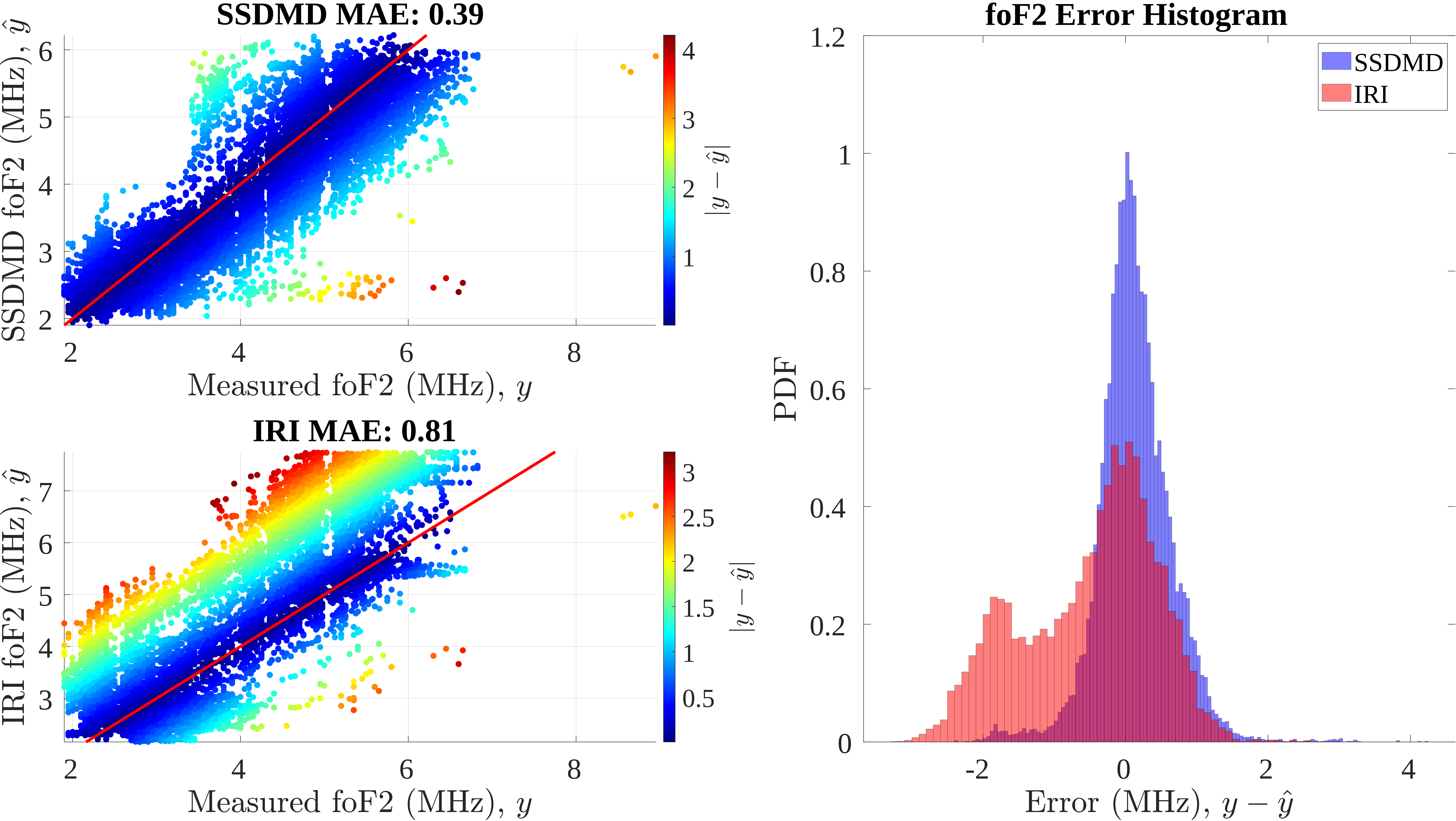

Hilbert Spectrum
Frequency (cycles/day)
Time (days)
Instantaneous Energy
0
20
40
60
80
100
120
0
2
4
6
8
10
12
2
4
6
8
10

Figure 12.



SSDMD MAE: 0.39
SSDMD foF2 (MHz), $\hat{y}$
Measured foF2 (MHz), y
$|y - \hat{y}|$
IRI MAE: 0.81
IRI foF2 (MHz), $\hat{y}$
Measured foF2 (MHz), y
$|y - \hat{y}|$
foF2 Error Histogram
SSDMD
IRI
PDF
Error (MHz), $y - \hat{y}$

Figure 13.

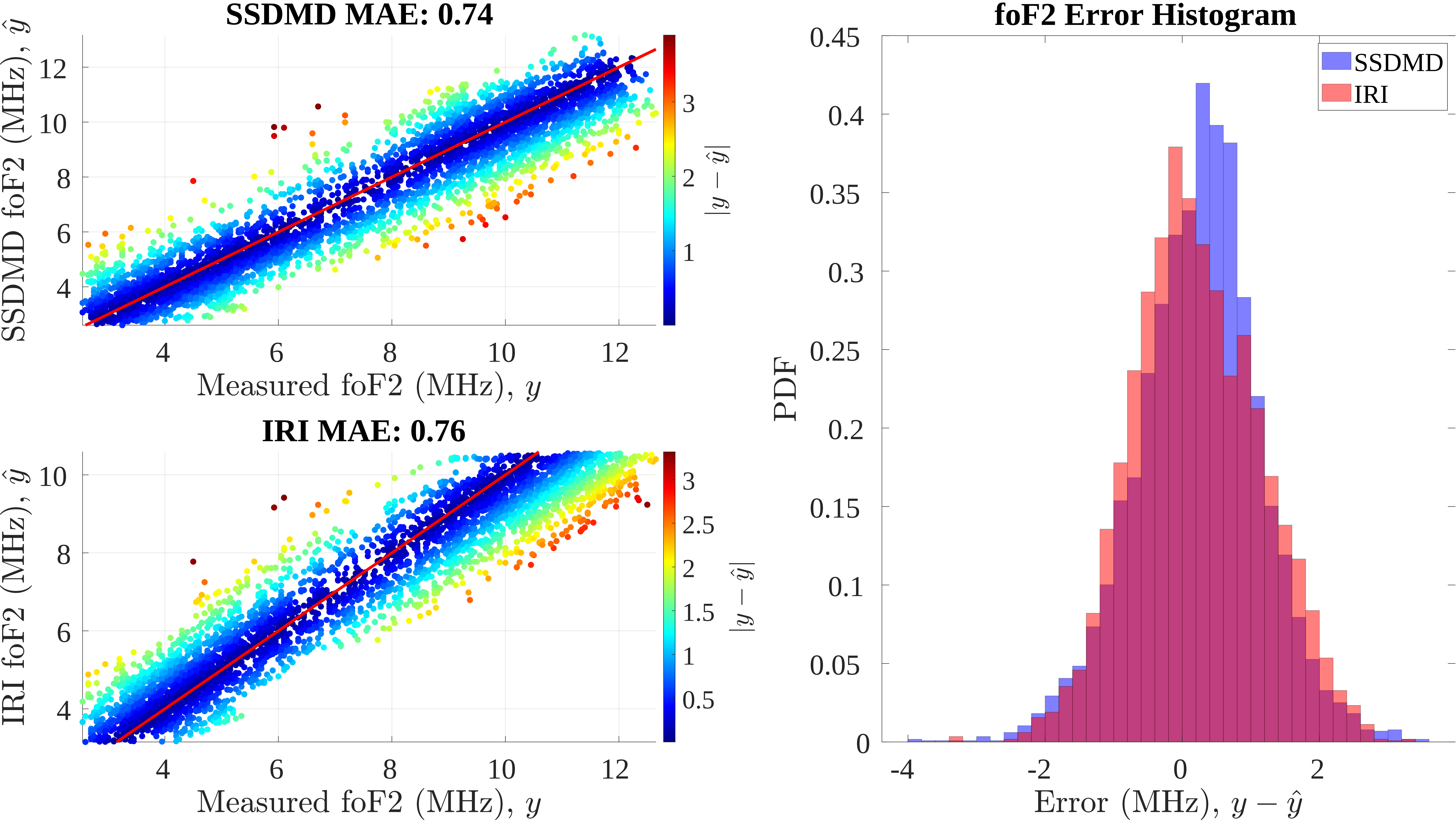
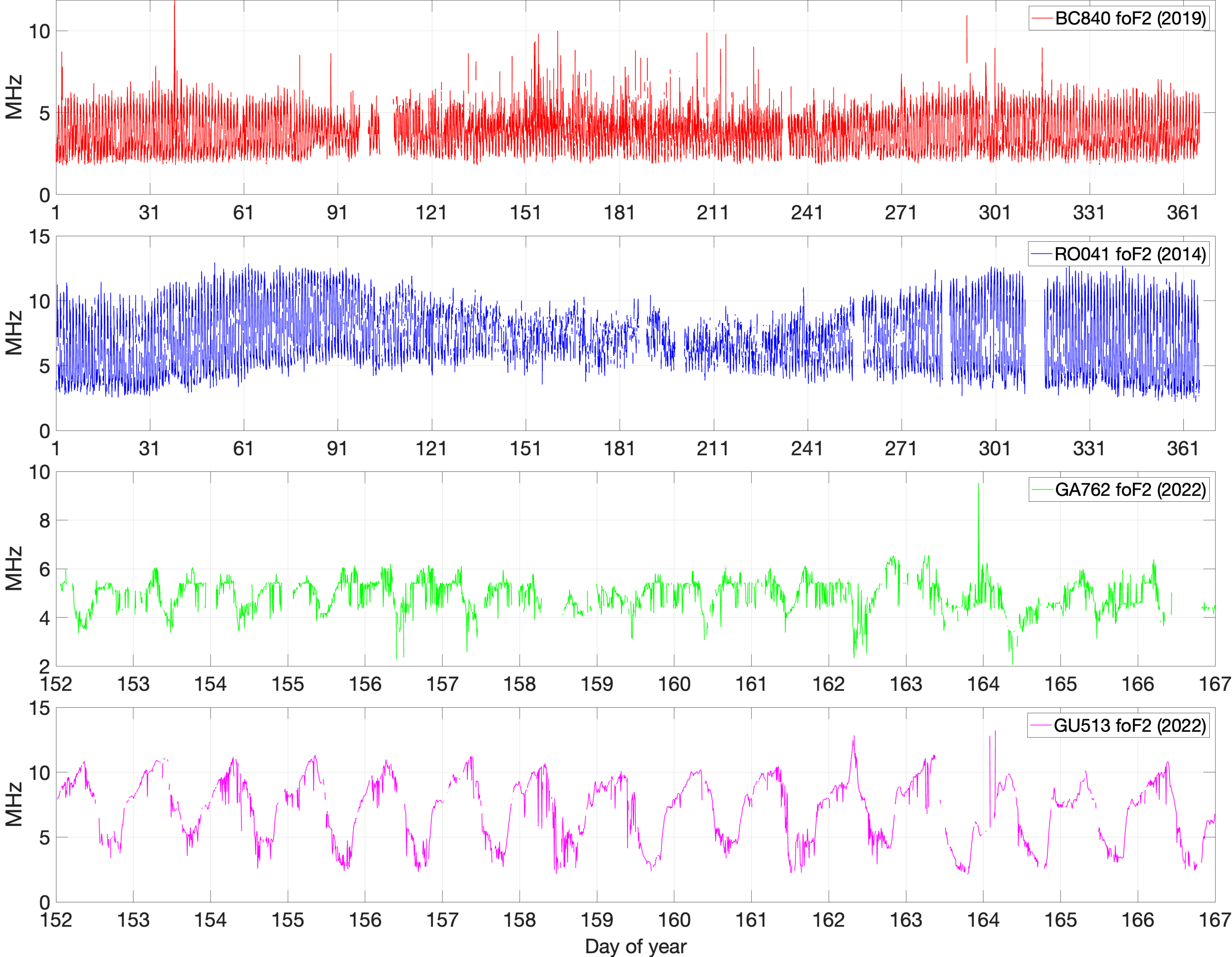

SSDMD MAE: 0.74
SSDMD foF2 (MHz), $\hat{y}$
Measured foF2 (MHz), y
$|y-\hat{y}|$
IRI MAE: 0.76
IRI foF2 (MHz), $\hat{y}$
Measured foF2 (MHz), y
$|y-\hat{y}|$
foF2 Error Histogram
SSDMD
IRI
PDF
Error (MHz), $y-\hat{y}$

Figure 8.

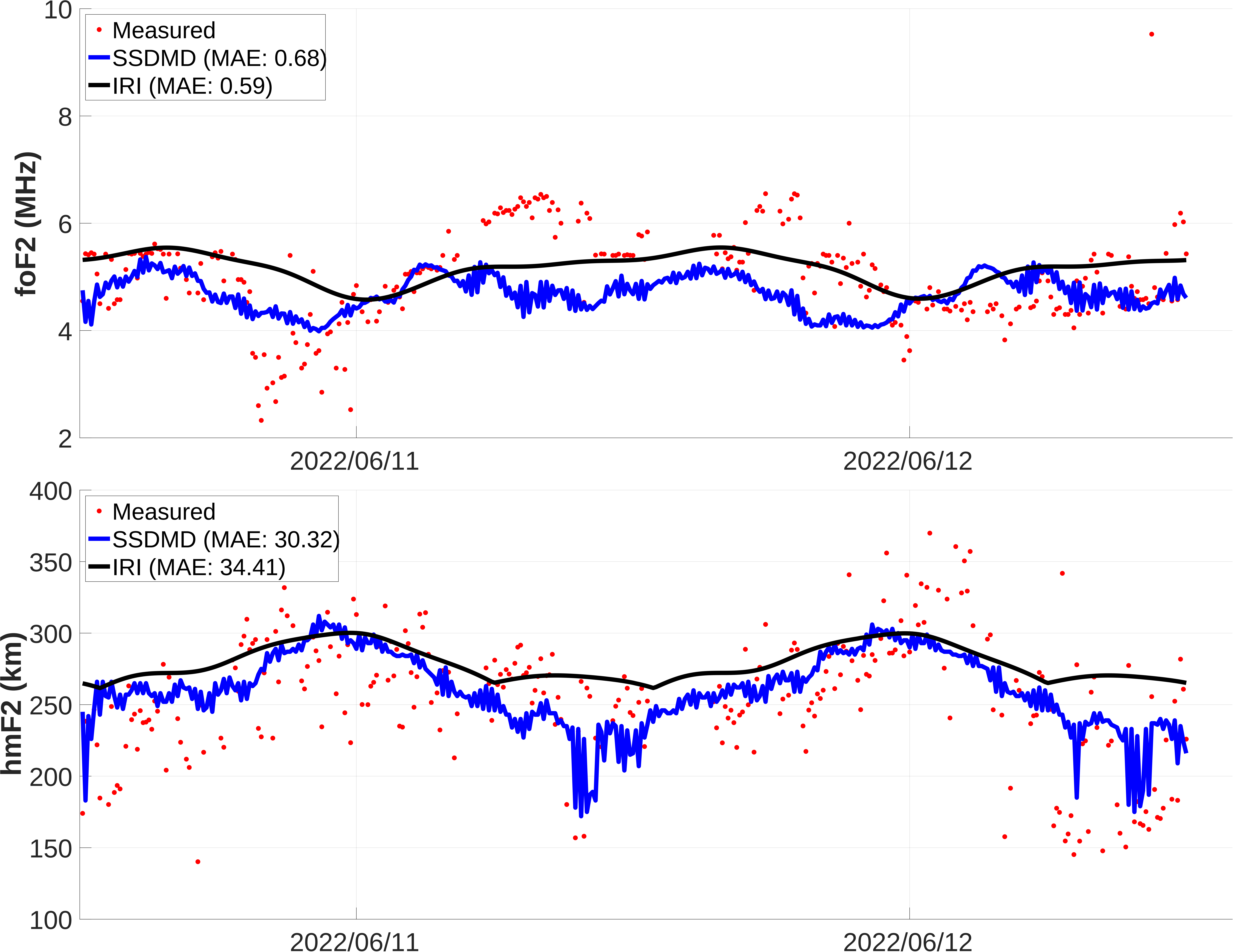


BC840 foF2 (2019)
RO041 foF2 (2014)
GA762 foF2 (2022)
GU513 foF2 (2022)
MHz
Day of year

Figure 10.

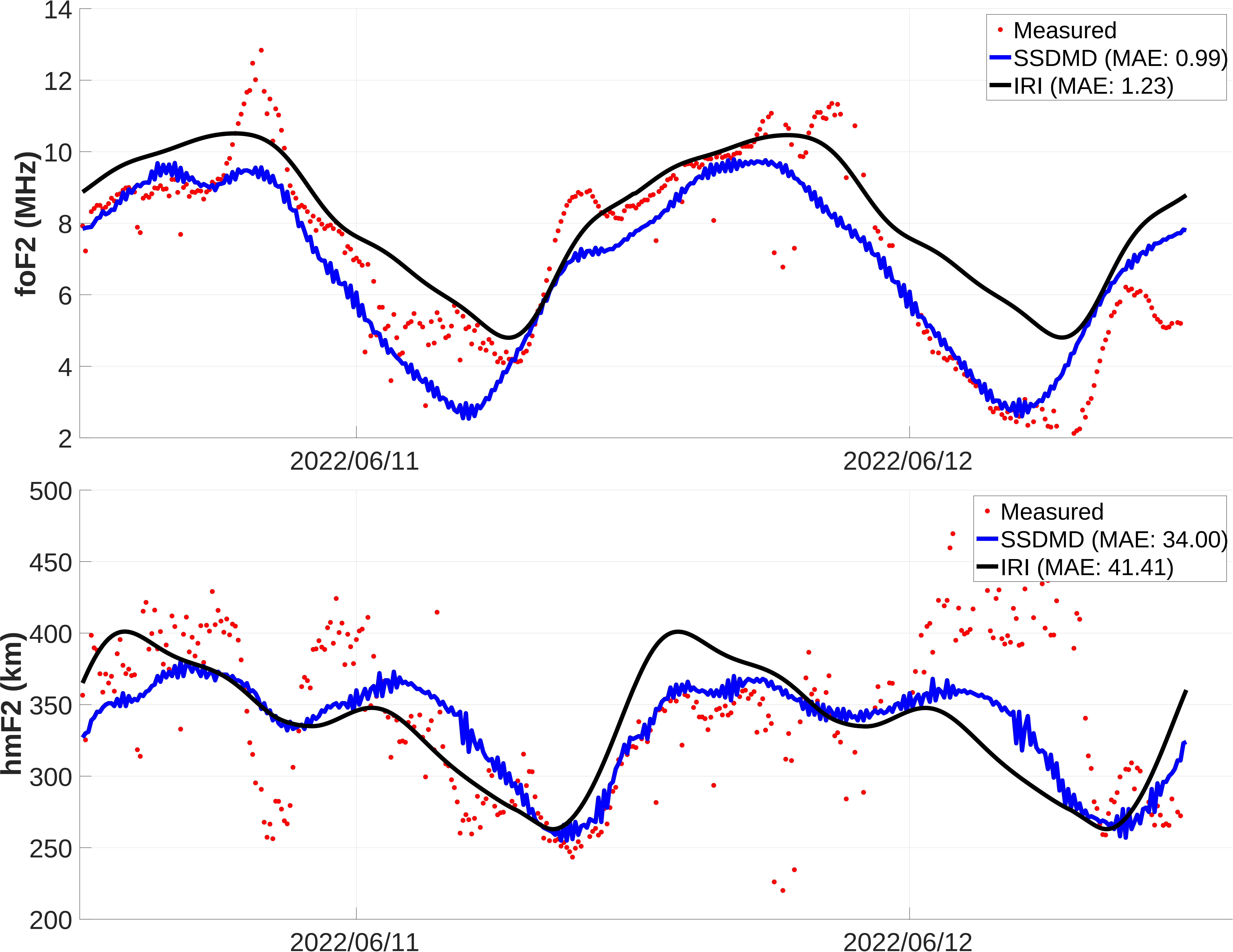
foF2 (MHz)
Measured
SSDMD (MAE: 0.33)
IRI (MAE: 0.44)
2019/03/05
2019/03/06
SSDMD (MAE: 0.39)
IRI (MAE: 0.91)
2019/12/22
2019/12/23
SSDMD (MAE: 0.54)
IRI (MAE: 0.86)
2019/10/20
2019/10/21
hmF2 (km)
SSDMD (MAE: 15.54)
IRI (MAE: 18.53)
SSDMD (MAE: 16.10)
IRI (MAE: 20.28)
SSDMD (MAE: 14.52)
IRI (MAE: 26.96)

Figure 16.



Measured
SSDMD (MAE: 0.68)
IRI (MAE: 0.59)
foF2 (MHz)
10
8
6
4
2
2022/06/11
2022/06/12
Measured
SSDMD (MAE: 30.32)
IRI (MAE: 34.41)
hmF2 (km)
400
350
300
250
200
150
100
2022/06/11
2022/06/12

Figure 17.



foF2 (MHz)
14
12
10
8
6
4
2
Measured
SSDMD (MAE: 0.99)
IRI (MAE: 1.23)
2022/06/11
2022/06/12
hmF2 (km)
500
450
400
350
300
250
200
Measured
SSDMD (MAE: 34.00)
IRI (MAE: 41.41)
2022/06/11
2022/06/12

Figure 11.

Measured
SSDMD (MAE: 0.70)
IRI (MAE: 0.96)
foF2 (MHz)
2014/10/29
2014/10/30
Measured
SSDMD (MAE: 0.73)
IRI (MAE: 0.86)
foF2 (MHz)
2014/12/07
2014/12/08
Measured
SSDMD (MAE: 0.56)
IRI (MAE: 0.76)
foF2 (MHz)
2014/12/07
2014/12/08
Measured
SSDMD (MAE: 16.91)
IRI (MAE: 11.58)
hmF2 (km)
2014/10/29
2014/10/30
Measured
SSDMD (MAE: 17.81)
IRI (MAE: 20.01)
hmF2 (km)
2014/12/07
2014/12/08
Measured
SSDMD (MAE: 18.43)
IRI (MAE: 20.20)
hmF2 (km)
2014/12/07
2014/12/08

Figure 14.

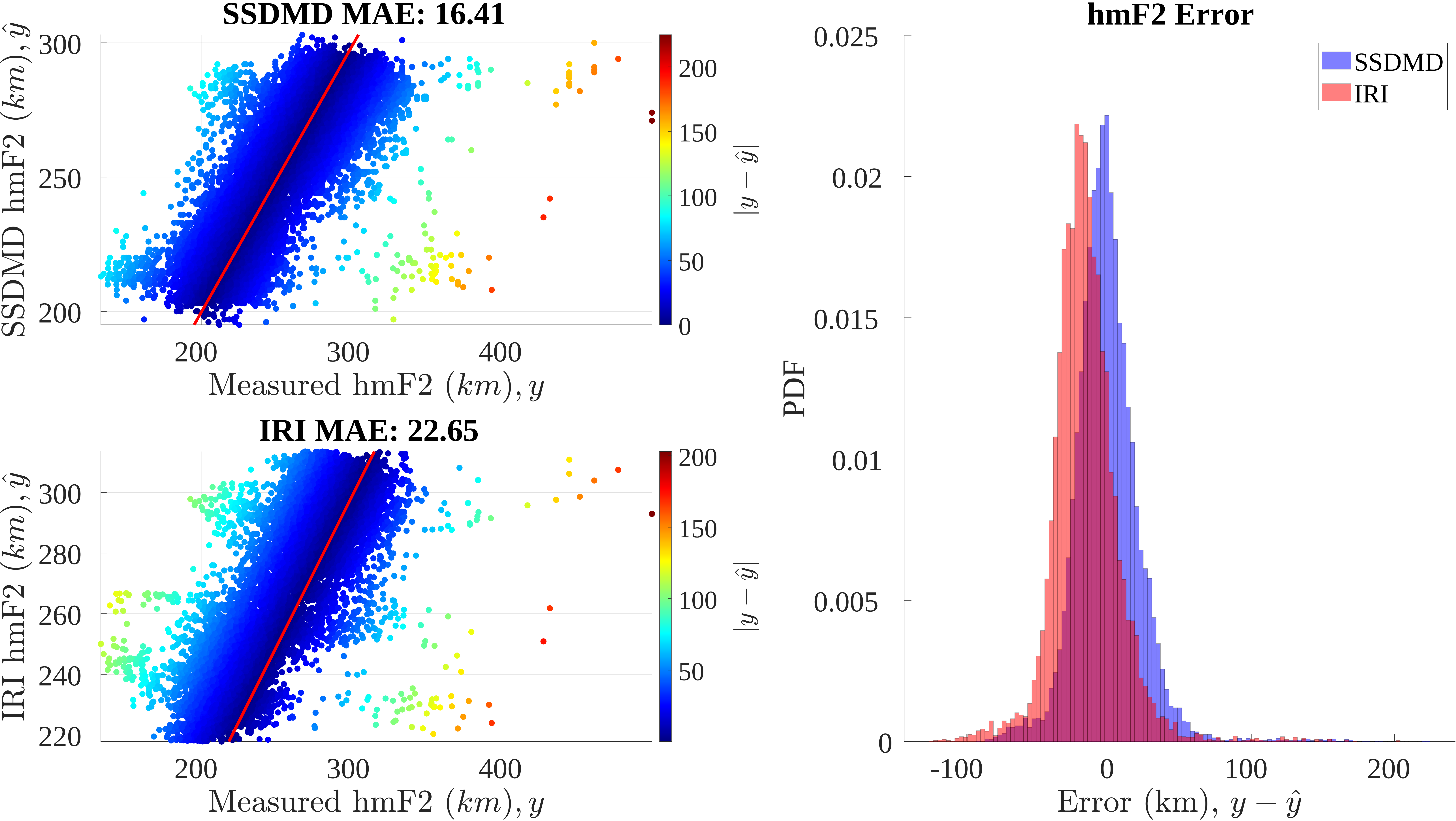
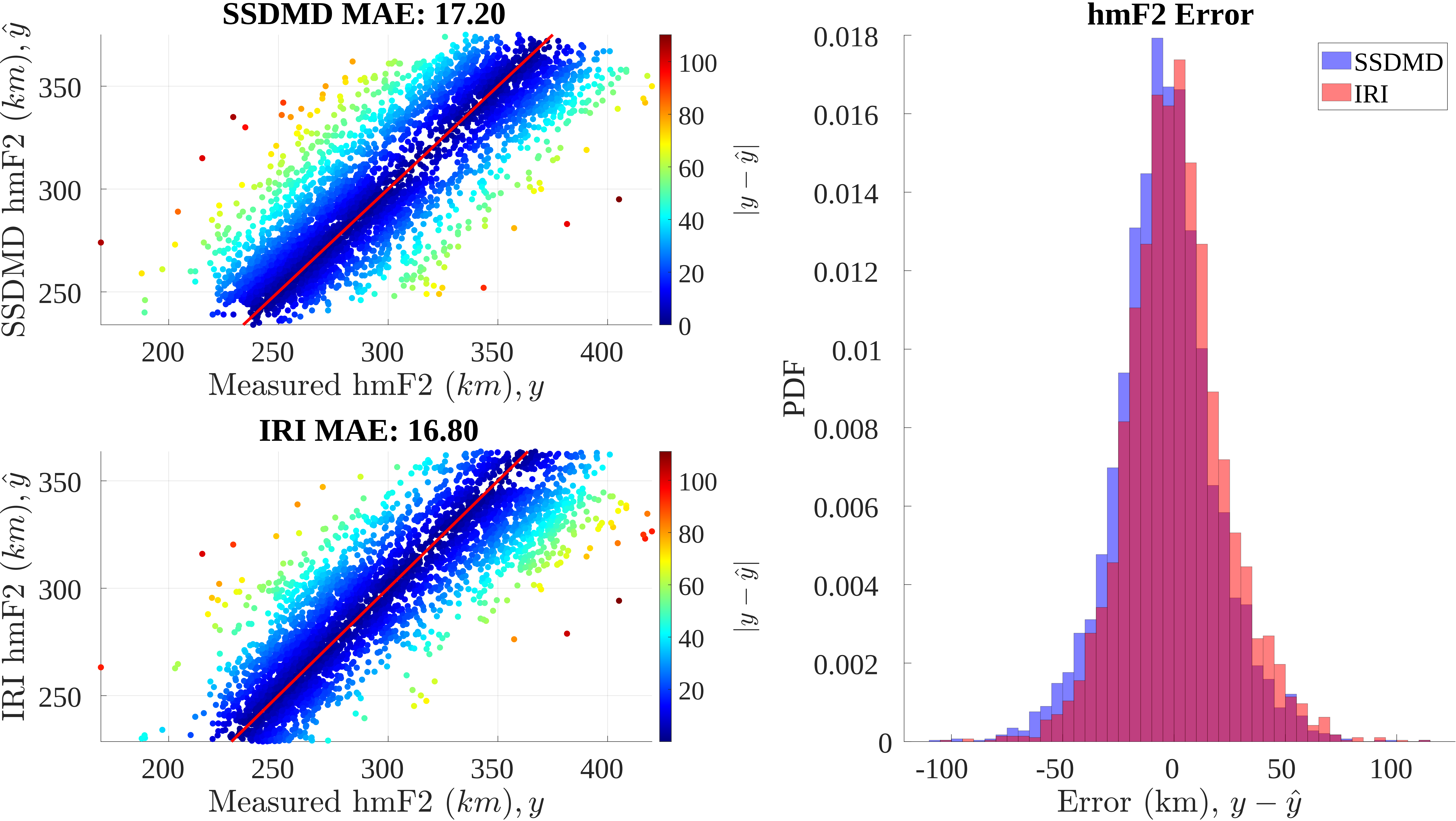

SSDMD MAE: 16.41
SSDMD hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
$|y - \hat{y}|$
IRI MAE: 22.65
IRI hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
$|y - \hat{y}|$
hmF2 Error
SSDMD
IRI
PDF
Error (km), $y - \hat{y}$

Figure 15.

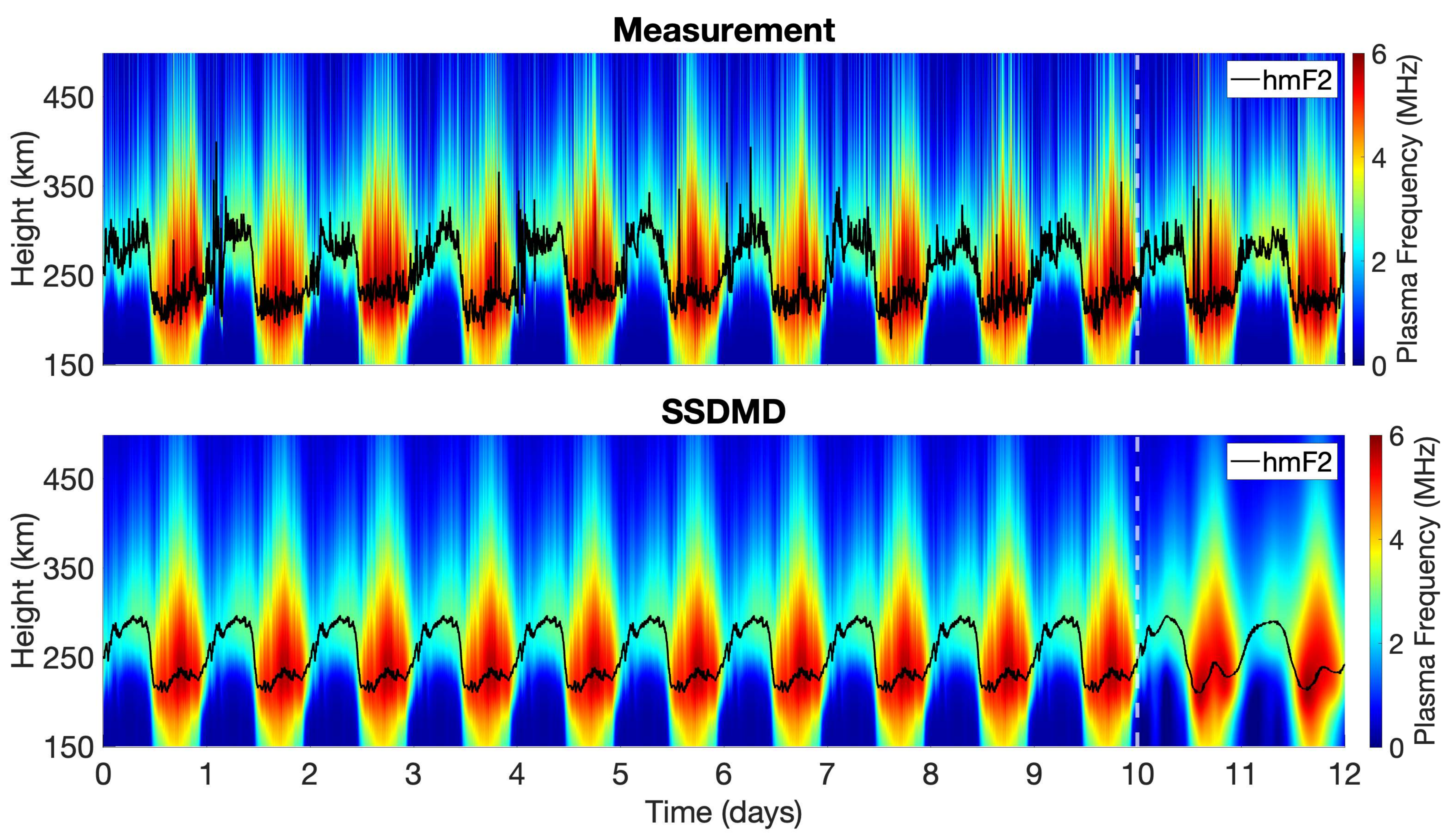


SSDMD MAE: 17.20
SSDMD hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
$|y-\hat{y}|$
IRI MAE: 16.80
IRI hmF2 $(km), \hat{y}$
Measured hmF2 $(km), y$
$|y-\hat{y}|$
hmF2 Error
SSDMD
IRI
PDF
Error (km), $y-\hat{y}$

Figure 9.


BC840 hmF2 (2019)
RO041 hmF2 (2014)
GA762 hmF2 (2022)
GU513 hmF2 (2022)
km
Day of year

Figure 7.



Measurement
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
6
4
2
0
SSDMD
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
6
4
2
0
0
1
2
3
4
5
6
7
8
9
10
11
12
Time (days)

Figure 18.

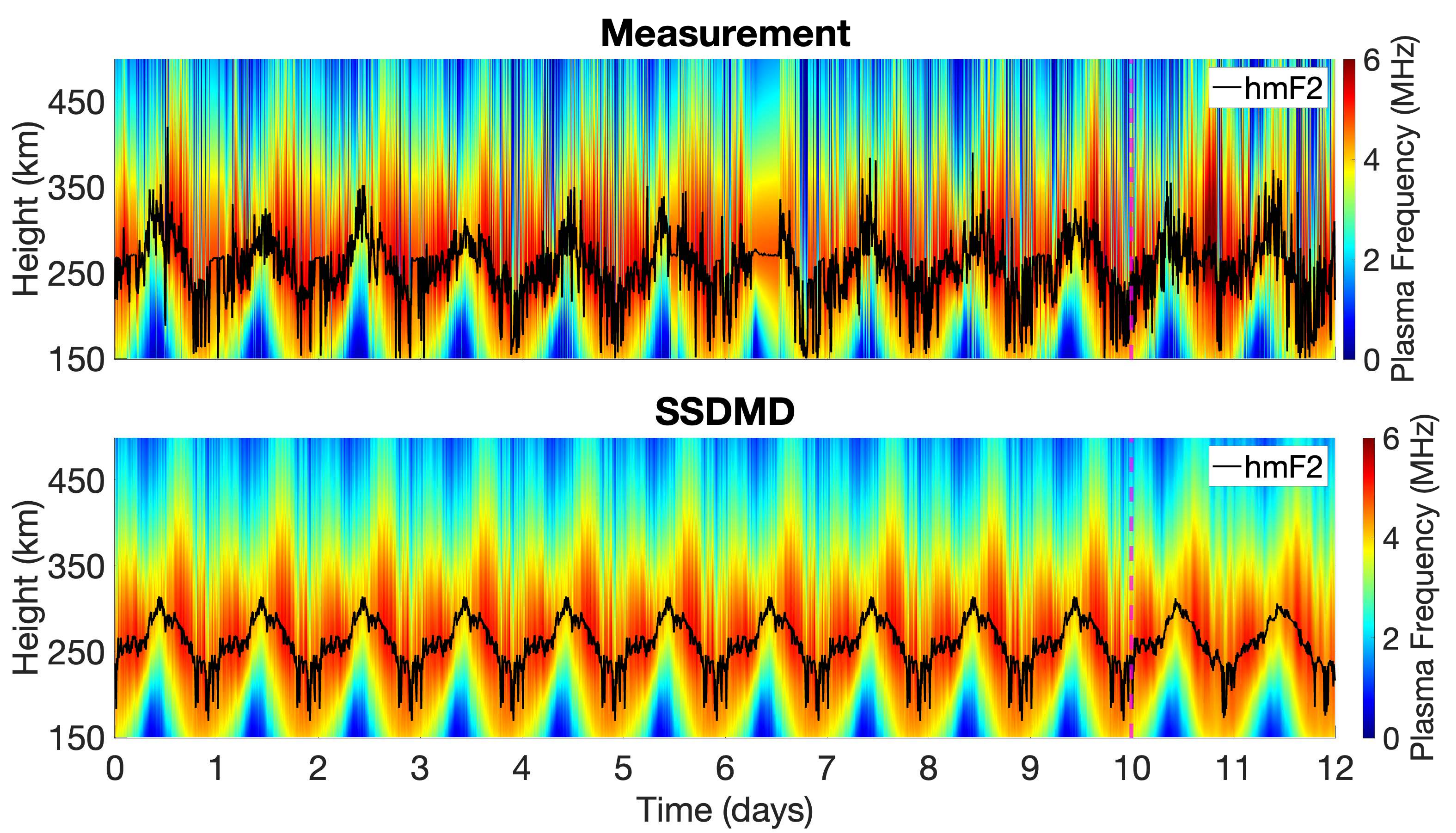
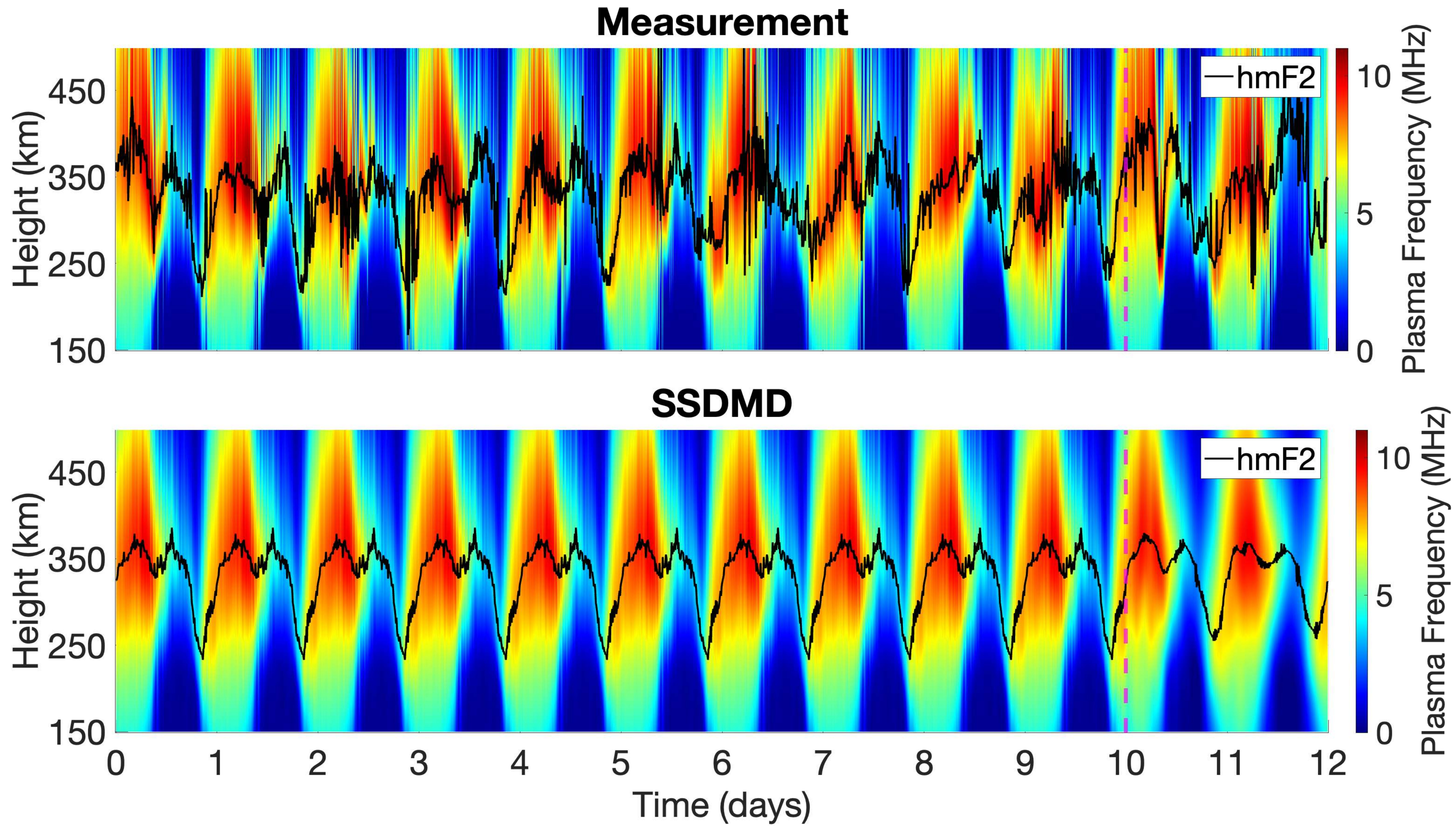
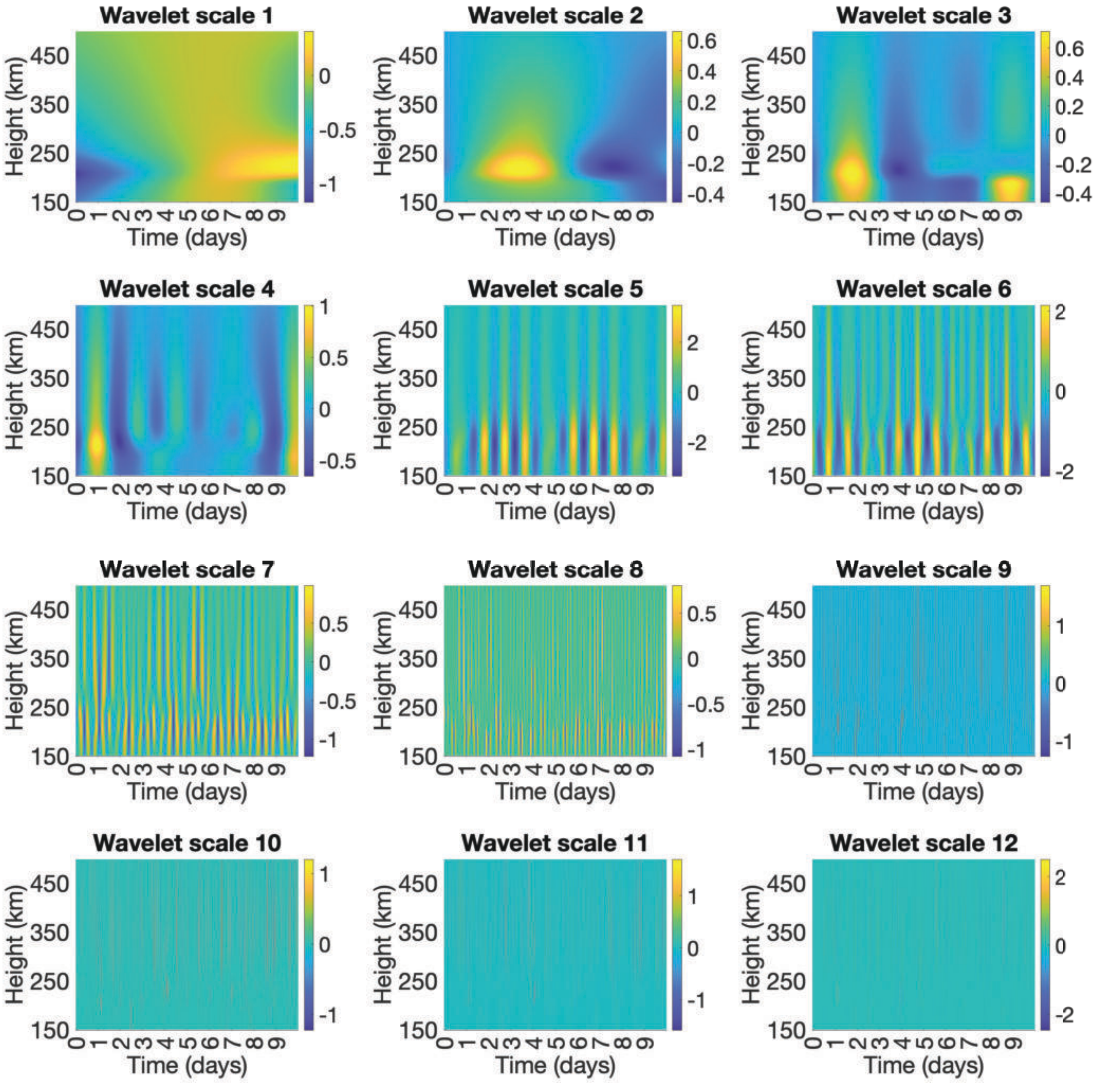

Measurement
hmF2
Height (km)
Plasma Frequency (MHz)
SSDMD
hmF2
Height (km)
Plasma Frequency (MHz)
Time (days)

Figure 19.



Measurement
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
10
5
0
SSDMD
hmF2
Height (km)
450
350
250
150
Plasma Frequency (MHz)
10
5
0
0
1
2
3
4
5
6
7
8
9
10
11
12
Time (days)

Figure 3.



Wavelet scale 1
Wavelet scale 2
Wavelet scale 3
Wavelet scale 4
Wavelet scale 5
Wavelet scale 6
Wavelet scale 7
Wavelet scale 8
Wavelet scale 9
Wavelet scale 10
Wavelet scale 11
Wavelet scale 12
Height (km)
Time (days)