Research article

Implementation of Adaptive-Bayesian DStoch technique for obtaining winds from MST radar covering higher altitudes

Manas Ranjan Padhy [a,*], Srinivasan Vigneshwari [a], M. Venkat Ratnam [b]

[a] *Sathyabama Institute of Science and Technology, Chennai, India*
[b] *National Atmospheric Research Laboratory, Gadanki, India*

A B S T R A C T

It is challenging to estimate winds accurately from higher altitudes using VHF-MST radar. The current study introduces the Adaptive-Bayesian Deterministic Stochastics Technique (ADStoch), which implements an Empirical Bayesian 1D prediction method using stochastics to analyze radar signals. A new and robust estimator for empirical wavelet shrinkage with Gaussian prior of the nonzero mean for wavelet coefficients is presented, which makes the current prior different from other priors. The mean parameters and the prior covariance hyperparameters follow a pseudo maximum likelihood method for computation. Details on the implemented algorithm developed from scratch using C# are also presented. This technique outperforms contemporary techniques discussed in this context that can recover signals buried in noise established based on the analysis of moment and quality. The estimated Wind is cross-validated for accuracy with the observed wind from the GPS radiosonde operated simultaneously. This technique can consistently extract 3D wind that can reach the range of 25.5 km–28.2 km, improving the conventional maximum altitude of 21.2 km in real time for the MST radar. It is concluded that the ADStoch analysis technique can effectively obtain VHF-MST radar signals at significantly higher altitudes, which is helpful in various scientific investigations.

1. Introduction

The VHF radars operating at around 50 MHz are potential tools to deliver three-dimensional winds essential for understanding various atmospheric dynamic processes. These are the only tools to date that can provide all the components of winds (zonal, meridional and vertical) with high temporal resolutions. The refractive index fluctuations observed in temperature and water vapour in the troposphere, temperature gradients from the lower stratosphere, electron density gradients, and turbulence from mesospheric altitudes within the inner and outer scales of turbulence primarily cause radar returns at these frequencies. Theoretically, these scales of turbulence extend up to 29.5 km in the lower atmosphere (Hocking, 1985) [1]. However, existing retrieval techniques, primarily relying on FFT, could yield moments and winds up to a typical range of 21.2 km because of the low signal-to-noise ratio in the lower stratosphere, where signals are frequently buried under noise. Obtaining information from these radars up to 30 km is useful particularly for understanding the process of energy exchange between stratosphere and troposphere (STE) (Das et al., 2020) [2], as part of background circulation and stratospheric influence on tropospheric weather and climate. Obtaining valuable signals from the upper troposphere (10–14 km) is also challenging from the retrieval techniques currently being used in the context of MST Radar

---

* Corresponding author.
  E-mail address: manas@narl.gov.in (M.R. Padhy).

(Ravindrababu et al., 2014) [3], as the refractive index is mainly due to temperature gradients alone, which are generally weak. Many authors have presented several techniques exploring the older configuration of MST radar (Rao et al., 1995), [4]. Recent architectural changes (Durga et al., 2020) [5] allow continuous mode observation using this radar. The older references predominantly based on signal processing techniques may not hold well now with evolving signal and noise characteristics provided by the new configuration of MST radar. Thus, there is a need to implement new analysis techniques to utilize radar-backscattered signals for various scientific investigations effectively.

Several data processing techniques earlier implemented on MST radar dataset for signal enhancement by noise suppression include the Bi-spectral process (Rao et al., 2008) [6], Principal component analysis (Rao et al., 2014) [7], Semi Parametric sparse Iterative Covariance-based Estimation (SPICE) technique (Eappen et al., 2015) [8], non-parametric and semi-parametric spectral techniques (Raju et al., 2019) [9]. These techniques could deliver a maximum altitude coverage of 20–21 km using the old MST dataset, which was discontinuous across incoherent integrations. The peak associating technique using Viterbi Data (VDA) (Bhatta et al., 2020) [10] has tried to predict peaks in higher ranges without carrying out signal enhancement. However, VDA could deliver a higher altitude of 21 km, similar to the radar's typical maximum standard range of 20–21 km. All these techniques have analyzed results based on only one data frame with SNR and noise parameters on an older MST radar dataset. The results need verification in terms of consistency, with several moments and quality parameters. Later, with the inclusion of the continuous mode of observation in MST radar during 2020, the signal enhancement with denoising using the adaptive ABlockJS technique was carried out by (Padhy et al., 2023). ABlockJS uses a non-Bayesian mixture model with a non-predictive scenario that could derive 3D Wind with moments and quality verified for consistency, which can cover a higher altitude range up to 25.2 km, thus improving the conventional maximum observable range of 20–21 km in real-time using MST Radar. The denoising techniques undermine the valid signal peaks in lower SNR scenarios in higher-altitude regions. This current study implements a Bayesian mixture model in a predictive scenario of parametric and non-parametric techniques that learn from data, eliminating the inconsistency that could arise with the exclusive use of parametric models. It establishes that the developed Adaptive-Bayesian DStoch (ADStoch) technique can effectively deliver winds from a much higher altitude range of up to 28.2 km from MST radar backscattered signals. This method compares results from other relevant techniques, where the signal enhancement could be possible, keeping a uniform scale of reference towards Moment and Quality parameters.

A model that does not contain any random component is known to be deterministic, which can be obtained by direct calculation. Initial conditions, parameter values, and relationships determine the deterministic model outputs. However, the stochastic model recognizes the random nature of the input components, accounting for the uncertainties caused by the varying behavioural characteristics and thus delivering distributions of relevant results for varying scenarios when run many times with the same input. Usually, stochastic models are more complex in computation. The results of stochastic models are difficult to interpret when compared to those of deterministic models.

The traditional technique for the MST radar data analysis uses FFT-based algorithms such as T_{FFT}. Recent studies have used Principal Component Analysis (PCA) and Iterative Adaptive Approach (IAA) (Uma et al., 2014, Raju et al., 2018) [7,11], group shrinkage overlapping Discrete Wavelet Transform (DWT) and wavelet-denoising (Babu et al., 2019, Thatiparthi et al., 2009) [12,13] covering lower altitudes up to 18 km using the old MST radar configuration. Also, the older analytic references may not hold well within a new set of signal and noise characteristics that have covered a single-cycle analysis, achieving relatively lower altitude coverage on the older MST radar, evaluating results mostly with SNR and noise with a single-cycle wind correlation. However, adaptive techniques are prone to fail due to a lack of adaptivity in a more extended dataset, raising inconsistency in estimating the moment and quality parameters and posing a challenge to establish the technique for efficient usage. Though the Bayesian approaches are relatively complex, they theoretically give an edge over non-Bayesian techniques. Meanwhile, in addition to FFT, several advanced methods have been implemented in this manuscript, such as the Hidden Markov Model (HMM), Covariance, Pseudospectrum, and YuleWalker, to find a suitable process that can work in real-time while covering higher altitudes. The set of 'Other Techniques' ($T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}, T_{FFT}$), implicates methods covered in the manuscript beside T_{ADS}, that constitute the set of 'All Techniques'.

The T_{HMM} encapsulates a Markov process with hidden states using the *1d*-HMM with Expectation Minimization (Crouse et al., 1998), [14]. The T_{CoV} technique uses an autoregressive estimate of *PSD*. The T_{PsS}, estimates the pseudo-spectrum from a signal using Schmidt's Eigen-space analysis method (Marple et al., 1989) [15], which uses a correlation matrix to estimate its frequency content, suitable for signals that are the sum of sinusoids with additive white Gaussian noise. The T_{YW}, uses autocorrelation to minimize the forward prediction errors in the least squares sense (Hayes, 2009), [16]. Though T_{YW} may slightly bias the frequency estimates of sinusoids in noise within applicable limits that do not affect the accuracy of estimating Wind. The T_{FFT}, calculates the power spectrum using FFT, following traditional MST Radar spectral analysis (Rao et al., 1995), [4]. The T_{ADS} is a Bayesian adaptive deterministic stochastic technique.

Standard wavelet has signal denoising application. Few studies have devised wavelet shrinkage methods that reconstruct signals from noisy data and recover anonymous information by a piecewise smoothing of noisy data (Donoho et al., 1995), [17]. The Bayesian Wavelet Shrinkage method is still an advanced approach that assumes a prior distribution with zero mean. Further, it is more realistic to use an estimator for the prior and follow the Empirical Bayes approach with the Bayesian shrinkage formulas to obtain more reliable results. Some of the previous studies have followed the Bayesian deterministic approach (Huang and Cressie, 2000) [18], and with independent prior (Chipman et al., 1997), [19]. Abramovich has used the squared error loss method with a loss function Bayesian approach rather than a shrinkage rule (Abramovich and Sapatinas, 1998), [20]. Donoho and others have reported that Decomp-Shrink is useable both for deterministic and stochastic signals, which is adaptive to the smoothness of the signal regardless of whether the signal has a sparse wavelet representation. An eBayes approach has been attempted to derive wind on a sparse noise peak condition that follows an empirical Bayesian model with Cauchy prior, using median thresholding rule, with a different posterior construct

without using *Stochastics* (Padhy et al., 2023), [21]. It may be noted that the assumed models, the priors (Gaussian vs. Cauchy), the procedure of estimating posterior, algorithm implementation and detailed software, pictorial demonstration of the process, thresholding rules, final results (posteriors), test dataset resolution and generation are quite different between the present study and earlier study [21].

The MST radar dataset has many nonzero wavelet coefficients in the higher-range range bins. Thus, the dataset's underlying signal in all these range bins can be assumed to contain deterministic components that are piecewise smooth and stochastic components with zero mean. Hence, it is necessary to use an adaptive mixed model scenario with a deterministic and a stochastic part for better noise suppression and, thus, signal enhancement. The implemented nonzero mean Gaussian prior towards the noisy signals can catch those signal components with both small or large wavelet coefficients based on a *prior covariance*. The implemented Bayesian-Adaptive DStoch technique T_{ADS}, is capable of recovering weak signals, particularly at higher altitudes, which *predicts* signals from noisy datasets optimally and efficiently.

2. The experiment

2.1. Description of experimental specification (base dataset)

The MST Radar was operated on April 29, 2022 for about 2.5 h. The specifications include a pulse width (PWD) of $16\mu s$, baud length (Baud) of $1\mu s$, and inter-pulse period (IPP) of $320\mu s$, giving a duty of 5.12%. The number of coherent integrations (NCI) and frequency sample points (nFFT) used were 256 and 256, respectively. Thus, the Doppler frequency and Doppler velocity range from -6.103 Hz to $+6.103$ Hz and -17.27 m/s to $+17.27$ m/s, respectively. The 2.7 km–29.5 km range, covered through 180 range-bins, has a 150 m resolution incorporating 5-non-coplanar beams (East, West, Zenith, North, and South), with 10 ^{o} tilt. The experiment has a frequency resolution of 0.047 Hz and a velocity resolution of 0.135 m/s. Thus, if ±2 frequency samples can be allowed in the analysis on the segmented dataset of 256 samples, the maximum expected error would be ±0.27 m/s. The dataset comprised 17 observation cycles. Each cycle has five beams, and every beam has five incoherent integrations in a segmented dataset of 256 IQ samples. It results in 425 frames, where each frame has 180 range-bins, and each range-bin has 256 IQ samples. We have also launched GPS radiosonde simultaneously during the above experiment to compare the winds with independent measurements. It has an accuracy of 0.5 m/s with a 5–6 m vertical resolution, later interpolated to 150 m to match the MST radar observations.

2.2. The preparation of data (collated dataset)

By averaging k independent random variables, the variance is reduced by a factor of k. A collated dataset prepared uses data across 5-incoherent integrations with 1280 ($256nFFT$ /frame x $5NCI$) sample points for each of these 5-beam. Every rage-bin comprising 1280 samples forms r segments of L samples each, which overlap, where $L \leq nFFT$. Due to measurement inaccuracy, the sinusoid amplitude in a frequency bin is often reflected in the adjacent frequency bin(s). With $L \leq nFFT$, the segment is zero-padded to $nFFT$ samples during periodogram calculation to improve the estimate. Thus, the dataset now becomes $x_r[n] = x[rR + n] * W[n], n = 0..L - 1$ and $r = 1,2..k$. This segmented data is further processed to eliminate bias and artefacts in the IQ data. A collated data sequence of 1280 samples across five incoherent integrations per beam has been created. The sequence is divided into 12 segments of 508 samples each. Each of these 12 segments is padded with four zeros to constitute a 512 IQ samples segment dataset x. Thus, the segmentation process improves the (frequency, velocity) resolution from (0.047 Hz, 0.135 m/s) to (0.0235 Hz, 0.067 m/s), also improves the estimation accuracy of

±2 frequency samples from ±0.27 m/s to ±0.135 m/s. The de-trending process eliminates DC bias in the segmented raw data. The data is then reshaped with a Kaiser window with optimized kβ of 5.85 to get the desired peak, main lobe width, and side lobe.

3. Methodology

In the following, we elaborate on the procedure to estimate the Normal mean using the *ADStoch* (T_{ADS}) while the methods for estimating PSD using the '*Other Techniques*' is elaborated in section *S.1* in the supplementary file. Similarly, section *S.2* elaborates on the moment and quality parameter while section *S.3* explains the methods for estimating Wind for '*All Techniques.*' (i.e. $T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}, T_{FFT}$). To derive the Normal mean using T_{ADS}, this technique starts with implementing the wavelet shrinkage method, which has been following three steps.

In the first step, the discrete wavelet transform transforms the data while yielding the empirical wavelet coefficients covered in Section *3.1*. In the second step, the empirical wavelet coefficients are shrunk toward zero. This process is evaluated to suppress noise based on the shrinkage rule explained in section *3.2*. The wavelet shrinkages are carried out by thresholding the wavelet coefficients. The thresholding process involves zero value replacement for the wavelet coefficients that have an absolute value below a prespecified threshold value. Finally, Section *3.3* transforms the processed empirical wavelet coefficients to the parent domain using the inverse wavelet transform in the third step.

3.1. Basic wavelet shrinkage

In the wavelet space, most energy tends to be concentrated in only a few coefficients. Suppose that we observe $Y(\bullet)$ at a discrete number $n = 2^J$ points; that is, we have data $Y = (Y_1, ..., Y_n)$, sampled uniformly in time, where $Yi = Y(t_i)$ and $t_i = i/n; i = 1, ..., n$.

The discrete wavelet transforms of Y is given by $w \equiv ((w^*_{J_0})', w'_{J_0}, .., w'_{J-1})' \equiv W_n Y$, where W_n is an orthogonal matrix and $w^*_{J_0} = (w^*_{J_0,0}, ..., w^*_{J_0,2^{J_0}-1})'$, which is a vector of scaling function coefficients and $w_j \equiv (w_{j,0}, ..., w_{j,2^{j}-1})'$ is a vector of wavelet coefficients at the jth scale $j \equiv J_0, ..., J-1$. The conventional wavelet-shrinkage approach transforms data using the discrete forward wavelet transform W_n, resulting in a decomposition. The data model incorporates measurement error, as in

$$w = \beta + \varepsilon, \qquad\qquad 1$$

where $w \equiv W_n Y$, signal $\beta \equiv W_n S$, which is not observed due to the noise contamination ε, and $\varepsilon \sim Gau(0, \sigma^2 I)$, where σ^2, represents noise variance.

Consider a model $(M.1)$ $\omega|\beta, \sigma^2 \sim Gau(\beta, \sigma^2 I)$ in our case, where the signal $\beta \equiv ((\beta^*_{J_0})', \beta'_{J_0}, .., \beta'_{J-1})'$ with an assumed prior distribution is represented as $(M.2)$ $\beta|\mu, \theta \sim Gau(\mu, \sum(\theta))$, where $\mu \equiv ((\mu^*_{J_0})', \mu'_{J_0}, .., \mu'_{J-1})'$ and suppose it has a deterministic mean structure where $\sum(\theta)$ is an $n X n$ covariance matrix structure constructed with user input θ, which describes the variability and the correlations in the signal. Initially, it can be assumed that. $(\mu'_{J_0}, ..., \mu'_{J-1})' \equiv 0$. The μ is the deterministic component to be computed using prior on the user input segmented dataset, represents the large-scale variation in β depicted in equation 1 makes,

$$\beta = \mu + \eta \qquad\qquad 2$$

where μ is the deterministic component, η is the stochastic component, and $\eta \sim Gau(0, \sum(\theta))$ represents the small-scale variation. Thus, the hyperparameters μ, σ^2, θ, requires to be dealt with in a Bayesian model.

3.2. Noise suppression using wavelet shrinkage

To suppress noise, the empirical wavelet coefficients w_j are "Shrunk" towards 0 based on a shrinkage rule. Usually, thresholding is done by replacing 0 for the values in w_j which are under a specified threshold value. The processed empirical wavelet coefficients are transformed to the parent domain using the inverse wavelet transform W'_n. In practice, the above $w^*_{J_0}$ scaling-function coefficients are not shrunk. The unknown noise parameter σ can be estimated using:

$$\widetilde{\sigma} = Median\ Absolute\ Deviation\ MAD\{\omega_{j-1,k}\} \equiv median\{|\omega_{j-1,k} - median\{\omega_{j-1,k}\}|\}/0.6745 \qquad (3)$$

The Bayesian wavelet shrinkage rule is obtained by specifying *aprior* for both β and σ^2 of equation 1.

3.3. Empirical Bayesian wavelet shrinkage

It is a procedure that requires a Bayesian hierarchical model as in **M.1** to estimate Noise Parameter σ, mean parameter μ and stochastic parameter η. Let $S = (S_1 ... S_n)$ be the sample space of one segment. Let the optimal predictor of β be $E(\beta, \omega)$, which we would like to transform back to the original dataspace. The inverse transform of W_n is W'_n, since W_n is an orthogonal matrix, equation 1 now becomes $W'_n w = W'_n \beta + W'_n \varepsilon$, because $W'_n \varepsilon$, is the white-noise measurement error in the data space $S \equiv W'_n \beta$, represents the signal that we need to predict. Due to the linearity, the optimal predictor of $S = (S_1 ... S_n)$ is $E(S|Y) = E(W'_n \beta | \omega) = W'_n E(\beta|\omega)$. Hence, after finding $E(\beta|\omega)$, the signal S can be optimally predicted. As in our case, we have assumed M.1 and M.2 as Gaussian, thus $\widehat{\beta}(\mu, \sigma^2, \theta) \equiv E(\beta|\omega)$. This is an empirical approach, and it offers improvements over the *apriori* specification of $(\mu'_{J_0}, ..., \mu'_{J-1})' \equiv 0$ and other methods for estimating σ^2.

Using deterministic and stochastic components, μ and η, allow it to recover a wider variety of signals by piecewise-smooth using a non-parametric regression model. In contrast, non-smooth signals often appear in time series. Nearby data tend to be more similar than distant data, making the stochastic description more acceptable than the deterministic description. Although it is possible to separate μ, η components asymptotically with some assumptions on μ and η from finite data, it is very difficult to distinguish between them.

Further, equation 2 can be achieved with high precision using an optimal predictor of β, such as $E(\beta|w) = \mu + \sum(\theta)(\sum(\theta) + \sigma^2 I)^{-1}(w - \mu)$. Based on estimates $\widehat{\sigma}^2$, an empirical Bayesian Decomp-Shrink rule is used, denoted by,

$$\widehat{\beta} = \widehat{\mu} + \sum(\widehat{\theta})\left(\sum\widehat{\theta} + \sigma^2 I\right)^{-1}(w - \widehat{\mu}). \qquad\qquad 4$$

Thus, the empirical Bayes predictor of $S \equiv W'_n \beta$ can be obtained using $\widehat{S} \equiv W'_n (\widehat{\mu} + \sum(\widehat{\theta})(\sum\widehat{\theta} + \sigma^2 I)^{-1}(w - \widehat{\mu}))$. For our dataset, hard thresholding fits well for a wavelet coefficient $w_{j,k}$ and a threshold level λ, which is represented by $T_H(\omega_{j,k}) = \omega_{j,k}$ if $|\omega_{j,k}| > \lambda$, '0' otherwise.

3.3.1. Estimation of noise parameter σ

The *MAD* estimator in equation 3 performs well whenever the signal is represented as a sparse wavelet by a piecewise smoothing of the deterministic signal. However, whenever the signal has stochastic components, the **MAD** estimator tends to overestimate the true value of σ. This overestimation of σ removes noise along with some essential and valid signal peaks. A new robust estimator for

estimating the noise variance based on Huang and Cressie [18] performs better than the Donoho-based estimator [17] for our dataset when used with Gaussian prior of the nonzero mean of Gaussian wavelet coefficients. This estimator

$$\widehat{\sigma}^2 \equiv (MAD\{Y_t\})^2, \tag{5}$$

is implemented for estimating σ, which is reliable when the signal β is either deterministic with $\sum(\theta) = 0$, or the signal β is *stochastics*.

3.3.2. Estimation of the deterministic mean parameter μ

The μ is not a user-defined parameter, which needs to be estimated. The Mean Parameter $\widehat{\mu}_{J_0}^*$ is estimated from $\widehat{w}_{J_0}^*$, which represents the large-scale features of the signal using

$$\widehat{\mu}_{J_0}^* = \widehat{w}_{J_0}^*, \tag{6}$$

assuming that its stochastic counterpart $\eta_{J_0}^* \equiv 0$, and declaring the scaling function coefficients $w_{J_0}^*$ to be purely deterministic. For the wavelet coefficients of the *j*th level, the deterministic trend μ_j could be considered as coming from components that are potential outliers in the normal probability plot of w_j. The slope of the fitted line for the normal probability plot is

$$\widehat{T}_j \equiv median\{|\omega_{j,k}|\}/0.6745. \tag{7}$$

For $k = 0, 1, \ldots, 2^j - 1$, let $q_{j,k}$ be the corresponding normal quantile of $\omega_{j,k}$ and let

$$r_{j,k} \equiv \omega_{j,k} - \widehat{T}_j q_{j,k} \tag{8}$$

be its residual. Thus, the estimate $\mu_{j,k}$ for $k = 0, 1, \ldots, 2^j - 1$ is given by

$$if \ |\omega_{j,k}| > \lambda_j \ then \ \widehat{\mu}_{j,k} = \frac{r_{j,k}^2}{\widehat{r}_j^2 + r_{j,k}^2} \ \omega_{j,k}, \ '0' \ otherwise, \tag{9}$$

where the threshold parameter λ_j is determined by

$$\lambda_j \equiv \widehat{T}_j \times max \ \{|q_{j,k}| : |\omega_{j,k}| < \widehat{T}_j |q_{j,k}|\}. \tag{10}$$

If $\{\omega_{j,0}, \omega_{j,1}, \ldots, \omega_{j,2^{j-1}}\}$ are normally distributed, either the threshold value λ_j becomes large, or the residuals $\{r_{j,0}, r_{j,1}, \ldots, r_{j,2^{j-1}}\}$ becomes small. Therefore, only a few points are estimated as deterministic trend components, but, importantly, these values are small. On the contrary, a signal with a large $\mu_{j,k}$ typically have a large $\omega_{j,k}$ and $r_{j,k}$, and the corresponding $\mu_{j,k}$ is hardly shrunk. Finally, the estimate of μ implemented as

$$\widehat{\mu} \equiv \left(\left(\widehat{\mu}_{J_0}^*\right)', \left(\widehat{\mu}_{J_0}^*\right)', \ldots, \left(\widehat{\mu}_{J-1}^*\right)'\right)', \text{for } j = J_0, \ldots, J-1, \widehat{\mu}_j \equiv (\mu_{j,0}, \mu_{j,1}, \ldots, \mu_{j,2^{j-1}})'. \tag{11}$$

3.3.3. Estimation of the stochastics component η

The Prior Covariance $\sum(\theta)$ is modelled with the corresponding stochastic component $\eta \equiv \left((\widehat{\eta}_{J_0}^*)', (\widehat{\eta}_{J_0}^*)', \ldots, (\widehat{\eta}_{J-1}^*)'\right)'$. The vector of θ is estimated by maximum likelihood based on the marginal distribution of the data w, with the values $\widehat{\sigma}^2$ and $\widehat{\mu}^2$ obtained from equations *5* and *11*, respectively. Using a zero-mean scale-independent model, the pseudo MLE of $(\theta \equiv \sigma_{J_0}^2, \ldots, \sigma_{J-1}^2)'$ is given by

$$\widehat{\sigma}_j^2 = max\left\{\frac{1}{2^j} \ (W_j - \widehat{\mu}_j)'(W_j - \widehat{\mu}_j) - \widehat{\sigma}^2, 0\right\}, \text{For } j = J_0, \ldots, J-1 \ . \tag{12}$$

The level-dependent wavelet shrinkage rule is developed based on equation *4*, for $j = J_0, \ldots, J-1, k = 0, 1, \ldots, 2^j - 1$ results into the following equation:

$$\widehat{\beta}_{j,k} = \widehat{\mu}_{j,k} + \frac{\widehat{\sigma}_j^2}{\widehat{\sigma}_j^2 + \widehat{\sigma}^2} (\omega_{j,k} - \widehat{\mu}_{j,k}), \tag{13}$$

w here $\widehat{\sigma}^2$ is given by equation *5*, $\{\widehat{\sigma}_j^2\}$ are given by equation *12* and $\{\widehat{\mu}_{j,k}\}$ are given by equation *9*.

3.4. Selection of model parameter(s)

For the *ADS* model, gaussian QMF, unlike Sym8 wavelet coefficients at *level 3* with level independent noise estimation using *Hard* threshold rule with enhanced adaptivity, is configured for the best performance. The Covariance and YuleWalker are configured with an order of 64. The PseudoSpectrum is configured with an order of 64 with a 128-length non-overlapping window. Finally, for *HMM*

model is configured without any parameter.

3.5. Application on segment dataset

Using the above models and segment data, the result obtained from the estimation process represents an independent segment of data, reducing variance by $1/k$ for averaging the k segments, each with the modified-*nFFT* length of 512 rather than *FFT* length of 256. Typically, non-overlapped segments ascertain maximum data independence. The bias reduces with increasing segment length ($L \leq$ *modified nFFT* of 512). Using more segments during segmentation relatively reduces the noise variance, which is further reduced by windowing each segment using the Kaiser window. Thus, the bias depends not exclusively on segment length and windowing but also on the selected technique among '*All Techniques*'.

3.6. Estimation of power spectral density (PSD), quality, moments and wind

The estimation of PSD using T_{ADS} is covered in section *3.1*. The estimation of PSD using '*Other Techniques*' ($T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}, T_{FFT}$) covered in Supplementary-Section-*S1a:S1e*. The method for estimating Moments, SNR, average band power, SINAD-ratio, Noise-Distortion content, SFDR, and Spur-power is elaborated in Supplementary-Section-*S2a: S2e*. The method for estimating *Wind* is elaborated in Supplementary-Section-*S3*.

4. Algorithm and software

Supplementary-Section *S.4* details the algorithm for estimating wind from the raw IQ comprising 19 steps that implement the complete mixture model to derive wind from the raw data using equations by equation *12:13* as depicted in sections *3.1 to 3.3*. The link to the implemented software source developed in C#.NET is made available in the code availability section. The pseudocode for implementing T_{ADS} from the segmented data is presented below.

Algorithm 1: DStoch
1.   Input: Segment data S_{sig}, wavelet parameters ($nlvl_3, QMF, LMax = 9$) for DStoch
2.   Calculate forward Wavelet Transform on S_{sig} and Wavelet Parameters using equation*1*.
3.   Calculate Adaptive-Bayesian DStoch Y_{Sig} using equation *2* using steps (3.1 to 3.6)
3.1. Estimate MAD (Median absolute deviation) using equation *3*
3.2. Using MAD, calculate Noise Variance σ^2 using robust estimator in equation *5*
3.3. Estimate DStoch mean parameter μ using equations *6*, *7*, *8*, *9*, *10*, *11*.
3.4. Calculate DStoch Structure yd for μ using IDWT
3.5. Calculate stochastic component η, using equations *12*, *13*.
3.6. Calculate DStoch Y_{Sig} while updating η by IDWT.
4.   Output: Store Y_{Sig}(deterministic and stochastic components) for current S_{sig}

5. Results and discussion

The base dataset depicted in section *2.1* generates a collated data sequence, as explained in section *2.2*. The dataset is analyzed separately using '*All Techniques*' for all beams and all cycles. The evaluation of outputs uses moments, quality, and wind, focusing on the highest altitude coverage that could be validated using GPS radiosonde data. The results show the measurement accuracy and the result consistency via 17 cycles of observations. The minor traceable signal differences, particularly in the higher ranges, may require good zoom on the figures covered in the results section.

5.1. Candidate Frame Analysis

The frame-normalized-*PSD* in Fig. 1 shows the presence of a signal at a higher range along with the variation in the signal strength,

which is a function of range and Doppler-frequency. The range-normalized-*PSD* depicts altitude coverage more prominently, allowing peak of individual range as a function of range and Doppler-frequency (demarcated by blue line). All the sub-figures represent maximum altitude coverage up to 25.5 km–28.2 km (established with correlation later), demarcated by a red dotted line. The supplementary *Fig-S1:S5* represents the same when analyzed using *'Other Techniques'*, which also depicts similar results with corresponding maximum altitude coverage of 25.2 km, 25.0 km, 22.9 km, 21.7 km, and 21.4 km, respectively. After that, the trace smears significantly up to 29.5 km. The RTI variation in T_{ADS} visually exhibits better SNR (Fig. 2), and is less noisy, especially noticeable at higher range. Similar characteristics observed in other beams are depicted in supplementary *Fig-S*6 : *S*9.

5.2. Moments analysis

The RTI variation of M_{FSNR} in East beam using T_{ADS} depicted in Fig. 3a1 exhibits higher FSNR than *'Other Techniques'* depicted vide *Fig-3a2:3a6.* Similarly, $M_{DopplerWidth}$ using T_{ADS} depicted in Fig. 3b1 exhibits consistency in a scale up to 1.6 Hz than using *'Other Techniques'* depicted vide Fig. 3b2:3b6. However, when compared with the FFT technique (T_{FFT}), a slight overestimate has been observed in $T_{HMM}, T_{ADS}, T_{PsS}, T_{CoV}, T_{YW}$ in sequence, especially in the lower range, where the SNR is usually high. The $M_{SignalPower}$ appears highest using T_{FFT} as depicted in Fig. 3c6 than T_{ADS}. However, in contrast, T_{FFT} exhibits poor noise performance in Fig. 3d6 than T_{ADS} depicted in Fig. 3d1, which appears pretty clean. As the signal power and noise cannot be separated, T_{ADS} can be treated better than *'Other Techniques'* due to the consistency, particularly in the higher altitudes demarcated by white dotted lines in Fig. 3. The supplementary-*Fig-S*10 : *S*13, represents similar results when analyzed for the other beams, making T_{ADS}, a better technique than all *'Other Techniques'*.

5.3. Quality Analysis

The supplementary Fig. S14 shows RTI variation of quality ($Q_{ABP}, Q_{SFDRR}, Q_{SP}, Q_{SINADR}, Q_{SINADN},$ and Q_{SINADP}) from the East Beams using *'All Techniques'*. Though, Q_{ABP} is observed to be highest while using T_{FFT} in Fig. S14a6, but it suffers significantly in Q_{SFDRR}, $Q_{SP}, Q_{SINADR}, Q_{SINADN}, Q_{QSINADP}$ in Fig. S14b6:4f6. The Q_{SFDRR} is observed to be best in T_{PsS} as it is a pseudo spectrum and noise approaches 0, unlike the spectrum in Fig. S14b4. The dynamic range is almost uniform in $T_{ADS}, T_{HMM}, T_{CoV}$ depicted vide Fig. S14b1:b3. The Q_{SP} is low and almost uniform, using T_{CoV}, T_{PsS}, T_{YW} depicted in Fig. S14c3:c5 but with less altitude coverage. The Q_{SP} is also uniform using T_{ADS}, T_{HMM} but covers higher altitudes as depicted in Fig. S14c1:c2. The evaluation of Q_{SINADR} shows best using T_{ADS} as depicted in Fig. S14d1. The Q_{SINADN} is evaluated to be uniform in all advanced techniques except in traditional techniques T_{FFT} as shown in Fig. S14e1:e5. The Q_{SINADP} is best among *'All Techniques'*, as depicted in Fig. S14f1. The supplementary Fig. S15:S18 represent similar results when analyzed for West, Zenith, North, and South beams. This RTI visual comparison reveals that the T_{ADS} performs better than *'Other Techniques'* for *'Quality'*.

5.4. Gain analysis

The average gain from Fig. 4 tabulated in Table. ST1, compares the quality obtained using T_{ADS} with the *'Other Techniques'* (T_{HMM}, T_{CoV}, T_{PsS}, T_{YW} and T_{FFT}), which is calculated for M_{FSNR} (5.07 dB, 3.14 dB, 3.04 dB, 1.91 dB, 3.87 dB), Q_{BP} (4.08 dB, 2.40 dB, 2.19 dB, 1.29 dB, 2.30 dB), Q_{NP} (−0.02 dB, −0.08 dB, −0.31 dB, −0.06 dB, −3.77 dB), Q_{SINADR} (5.07dBc, 3.14dBc, 3.04dBc, 1.91dBc, 3.87dBc), Q_{SFDRR} (0.27dBc, 0.56dBc, 1.75dBc, 0.71dBc, 1.78dBc), Q_{SINADP} (−1.00 dB, −0.94 dB, −1.21 dB, −0.94 dB, −4.77 dB), Q_{SP} (−1.16 dB, −1.19 dB, −1.95 dB, −1.26 dB, −8.27 dB). The supplementary Fig. S19 : S22, represents similar results when analyzed towards the rest of the beams, making T_{ADS}, a better technique as it quantifies with better average gain than the *'Other Techniques'*. The comparison quantifies better FSNR, relatively more band power, less Noise power, better SINAD ratio, good SFDR, lowest SINAD distortion power,
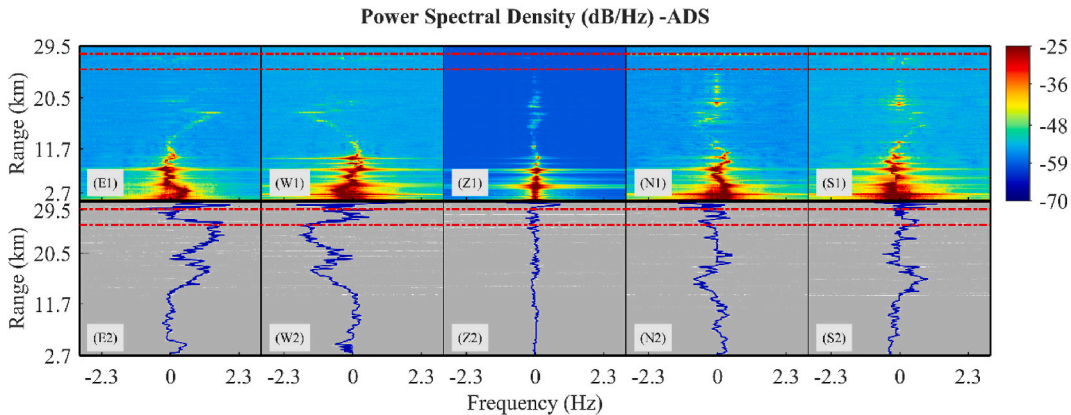

Fig. 1. (E1,W1,Z1,N1,S1) and (E2,W2,Z2,N2,S2) represent the frame and the range normalized PSD using Adaptive Deterministic Stochastic technique (T_{ADS}), for *'All Beams'* (East, West, Zenith, North and South) of the 5th cycle on Apr. 29, 2022 during 21:59:35 IST to 22:07:58 IST.
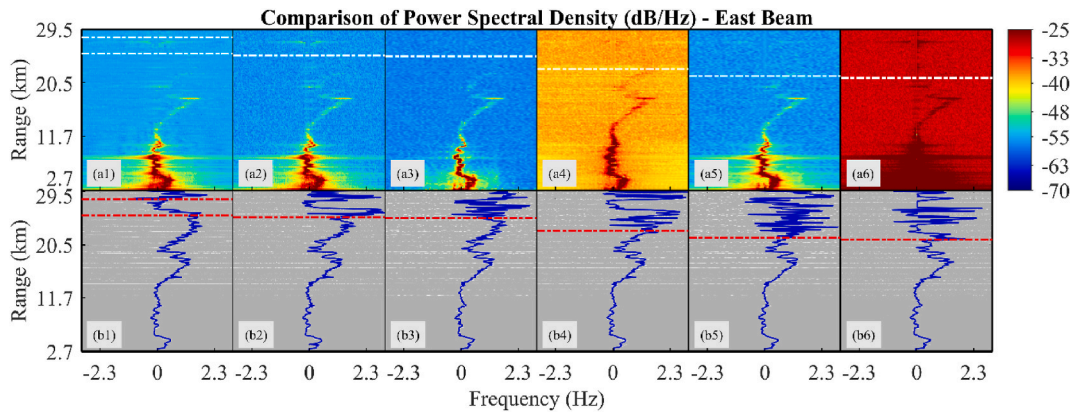
Fig. 2. (a1:a6), (b1:b6) represent the frame and the range normalized PSD using 'All Techniques' ($T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}$ and T_{FFT}), respectively, for the East Beam of the 5th cycle on Apr. 29, 2022 during 21:59:35 IST to 22:00:59 IST.
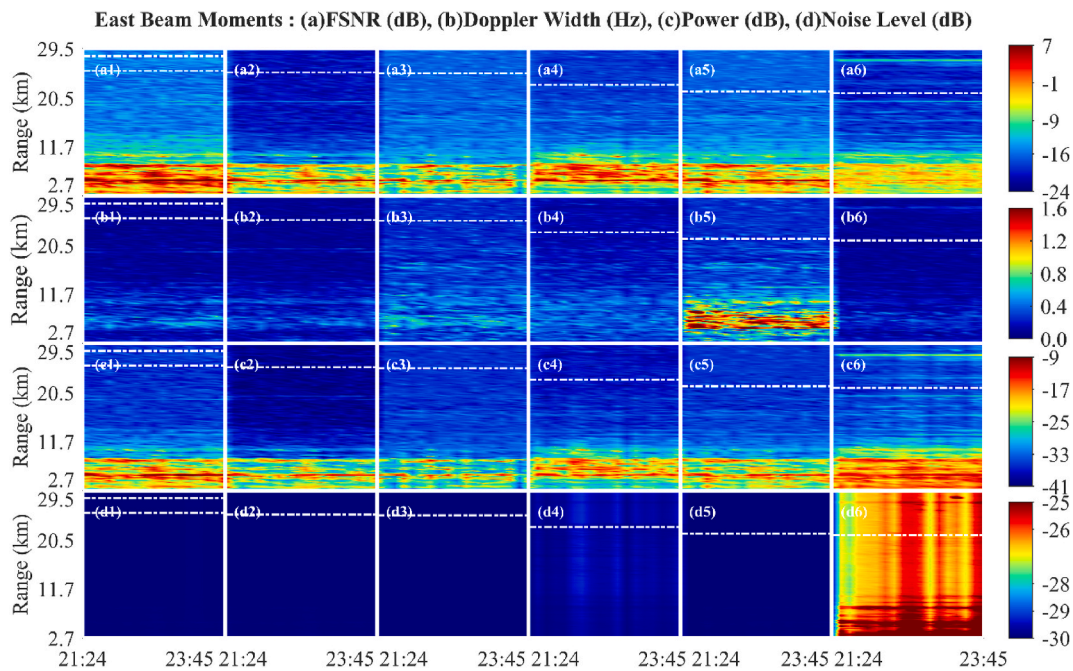

Fig. 3. (a1:a6), (b1:b6), (c1:c6), (d1:d6) represent the Fundamental Signal to Noise Ratio (M_{FSNR}), Doppler-width $M_{DopplerWidth}$, Signal Power $M_{SignalPower,}$ and Noise-level $M_{NoiseLevel}$ using 'All Techniques' ($T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}$ and T_{FFT}), respectively, for the East Beams of 1:17 cycles on Apr. 29, 2022 during 21:24:37 IST to 23:45:50 IST.

and lowest Spurious power.

5.5. Wind analysis

Fig. S23 shows the RTI variation of 'Wind' ($U_{Zonal}, V_{Meridional}, W_{Vertical}$), WS_{HW} (Horizontal Wind Speed) and WD_{HW} (Wind Direction) for 17-cycles derived using 'All Techniques'. The 'Wind', WS_{HW}, WD_{HW}, has consistent variation below the maximum altitude (demarcated with white dotted lines), above which the estimation is inconsistent. *Fig. 5 compares the mean of 17-cycles in WS_{HW}, $U_{Zonal}, V_{Meridional}$ derived using 'All Techniques' with that of the $WS_{HWGPS}, U_{ZonalGPS}, V_{MeridionalGPS}$. The marginal root mean square difference observed between the MST wind and the GPS Wind, demarcated in red, resulted from the variation in the Wind during 2.5 h of observation.*

5.6. Wind validation

Fig. 6a1:6a3 shows the mean correlation of $WS_{HW}, U_{Zonal}, V_{Meridional}$, derived using T_{ADS}, with that of the WS_{HWGPS} from 2.7 km to
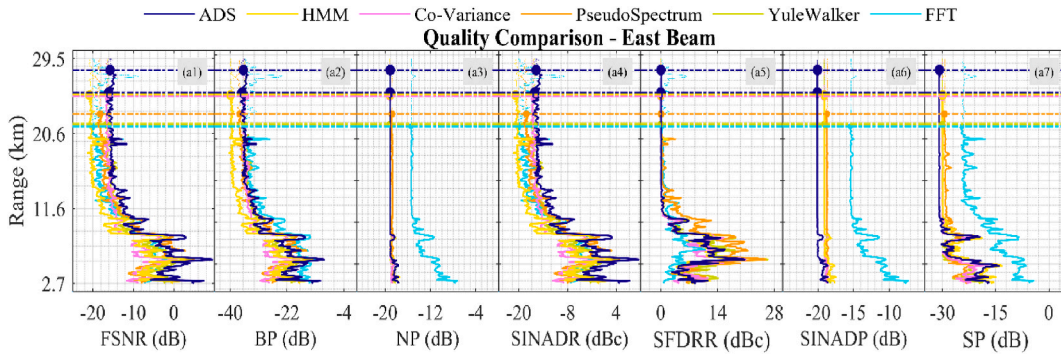
Fig. 4. (a1:a7) represent the mean variation in the 1:17 cycles of the Fundamental Signal to Noise Ratio (M_{FSNR}), Band Power (Q_{BP}), Noise Power (Q_{NP}), Signal to Noise and Distortion Ratio (Q_{SINADR}), Spurious Free Dynamic Range Ratio (Q_{SFDRR}), Signal to Noise and Distortion Power (Q_{SINADP}), and Spurious Power (Q_{SP}) using '*All Techniques*' ($T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}$ and T_{FFT}), respectively for the East Beams on Apr. 29, 2022. during 21:24:37IST to 23:45:50IST.
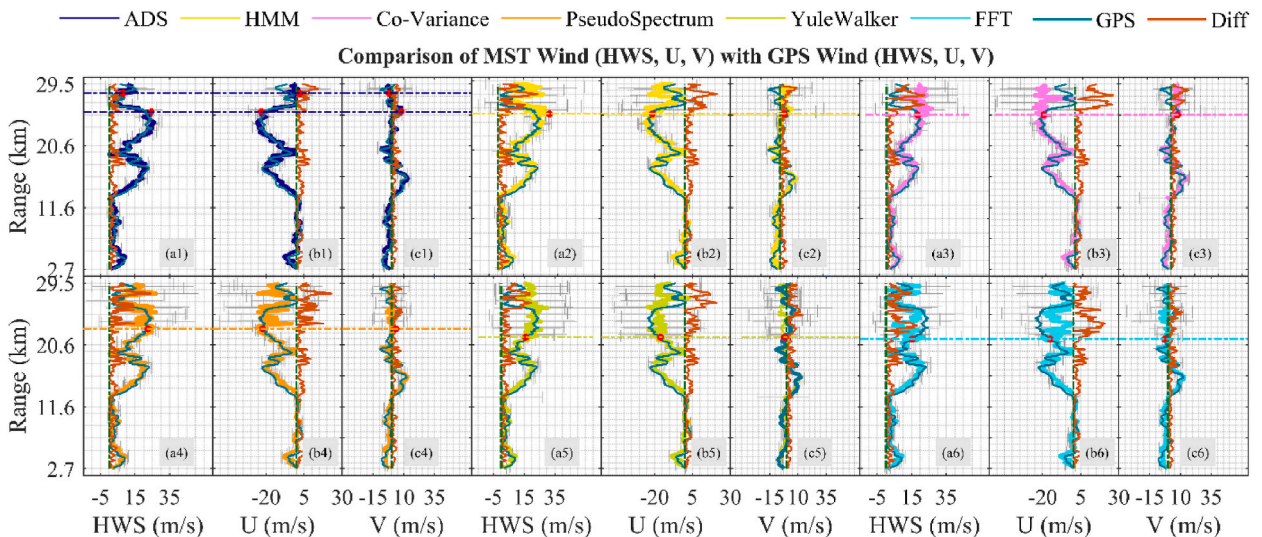

Fig. 5. (a1:a6), (b1:b6), (c1:c6) represent the WS_{HW} (Horizontal Wind Speed), $U_{Zonal}, V_{Meridional}$ using '*All Techniques*' ($T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW}$ and T_{FFT}), respectively for 1:17 cycles on Apr. 29, 2022 during 21:24:37IST to 23:52:49IST.

25.5 km with R = 0.95 from 25.5 km to 28.2 km with R = 0.79, from 28.2 km to 29.5 km with R = −0.01, $U_{ZonalGPS}$ from 2.7 km to 25.5 km with R = 0.96, from 25.5 km to 28.2 km with R = 0.85, from 28.2 km to 29.5 km with R = −0.53 and $V_{MeridionalGPS}$ from 2.7 km to 25.5 km with R = 0.87, from 25.5 km to 28.2 km with R = 0.68, from 28.2 km to 29.5 km with R = −0.40 shows good measurement accuracy for the estimated wind while using T_{ADS}. Similarly, the supplementary Fig. S24 depicts the correlation of wind derived using '*All Techniques*' for the possible altitudes. Fig. 6b1:6b3 shows the cumulative correlation of $WS_{HW}, U_{Zonal}, V_{Meridional}$ ([2.7 km–17.5 km]: [2.7 km–29.5 km]) derived using T_{ADS}, with that of the $WS_{HWGPS}, U_{ZonalGPS}, V_{MeridionalGPS}$. If the cutoff for $WS_{HW}, U_{Zonal}, V_{Meridional}$, can be assumed to be 0.9, 0.9, 0.85 (demarcated with red line), then the maximum altitude shows a sharp fall in coefficient of correlation beyond maximum altitude in this context corresponding to '*All Techniques*', i.e. (25.5 km:28.2 km, 25.2 km, 25.0 km, 22.9 km, 21.7 km, 21.4 km). The blue-dashed line in Fig. 6b1:6b3 up to an altitude of 25.5 km has a good correlation, and the blue-solid line at 28.2 km with a reduction in correlation shows that the T_{ADS}, can be better trusted to range coverage up to 25.5 km, beyond which it depends upon the need for measurement accuracy. Beyond 28.2 km T_{ADS} completely fails with a negative correlation. This confirms better performance of ADS over '*Other Techniques*'.

5.7. Analytical study of T_{ADS} technique

An analytical study has been carried out that compares and demonstrates the results of one sample segment, which are analyzed using the ADS technique with Huang and Donoho-based noise variance estimators while using Gaussian prior with Gaussian wavelet coefficients annexed in the Supplementary Analysis Section SAS (*SAS 1*: *SAS 9*).
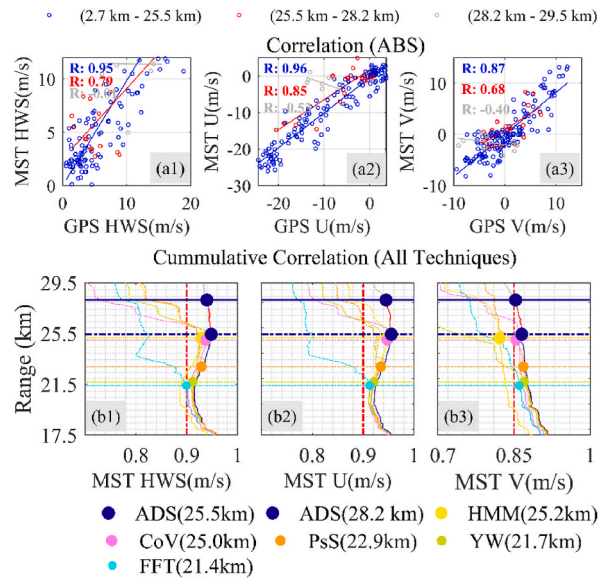
Fig. 6. (a1:a3) represents the correlation of MST Wind (WS_{HW}, U_{Zonal}, $V_{Meridional}$) with that of GPS Wind (WS_{HWGPS} $U_{ZonalGPS}$, $V_{MeridionalGPS}$) respectively, using Adaptive Deterministic technique from (2.7 km–25.5 km, 25.5 km–28.2 km, 28.2 km–29.5 km). (b1:b3) represent the cumulative correlation (from 2.7 km to 29.5 km) for the Mean (1:17 cycles) of MST Wind (HWS, U, V) with that of GPS Wind (HWS, U, V), respectively, using *'All Techniques'* (T_{ADS}, T_{HMM}, T_{CoV}, T_{PsS}, T_{YW} and T_{FFT}).

6. Summary and conclusions

The current study is the first to develop and use the Adaptive-Bayesian DStoch (ADStoch) technique with a predictability-based hybrid model to derive high-resolution 3D-wind components from MST radar raw IQ data. The manuscript details this approach, including its advantages over five *'Other Techniques'* and its distinctiveness. The key findings of this investigation are summarized below.

1. It employs a novel collating concept based on incoherent integration, leveraging the continuous observation mode of the new MST radar, which was impossible in the earlier configuration of MST radar. The introduced segmentation concept enhances signal strength even further.
2. The technique is well evaluated based on various signal and noise Quality parameters such as average band power, SFDR, spurious power, SINAD, SNR, distortion power, and full band noise power.
3. ADStoch outperformed *'Other Techniques'* at higher altitudes, as evidenced by Candidate Frame Analysis, Moment Analysis, and Quality Analysis. The Gain Analysis reveals a gain of 5dB–8dB using ADStoch, which is quite significant, especially at higher altitudes. The accuracy of wind computed using ADStoch is relatively high, as demonstrated by a comparison with independent GPS radiosonde winds.
4. The study demonstrates and concludes that the Adaptive-Bayesian DStoch (ADStoch) approach can reliably derive 3D-wind, covering a minimum higher altitude of 25.5 km, depending on the demand for measurement accuracy up to 28.2 km, which enhances the usual new MST radar's conventional maximum altitude coverage up to 21.2 km in real-time. It is highly beneficial to various scientific research using VHF radars.

However, there are a few limitations to using this technique in real-time. One needs a good computing facility, as we have currently done in our case. The inter-cycle delay is sufficient to carry out the predictive Bayesian data analysis before the completion of the current data cycle. As the data is collected onboard in digital receivers in a double buffer, the analysis in the computer is treated in real-time for the current data collection.

Code availability

The related data and code are available for anonymous downloading at the Datacenter public repository Link: https://www.narl.gov.in/datacenter/BayesianADStoch comprising the color versions of the figures and the supplementary figures in this article and data in. nc format. The related code is available at the public repository https://github.com/manasnarl/BayesianADStoch/.

CRediT authorship contribution statement

Manas Ranjan Padhy: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Srinivasan Vigneshwari:** Writing – review & editing, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.heliyon.2024.e26316.

References

[1] W.K. Hocking, Measurement of turbulent energy dissipation rates in the middle atmosphere by radar techniques: a review, Radio Sci. 20 (6) (1985) 1403–1422.
[2] S.S. Das, K.V. Suneeth, M. Venkat Ratnam, K.N. Uma, M.D. Rao, A. Narendra Babu, Long-term observations of stratosphere-troposphere exchange using MST Radar and Aura MLS measurements over a tropical station Gadanki, Radio Sci. 55 (2020) e2019RS006969, https://doi.org/10.1029/2019RS006969.
[3] S. Ravindrababu, M. Venkat Ratnam, S.V. Sunilkumar, K. Parameswaran, B.V. Krishna Murthy, Detection of tropopause altitude using Indian MST radar data and comparison with simultaneous radiosonde observations, J. Atmos. Sol. Terr. Phys. (2014) 121240–121247.
[4] P.B. Rao, A.R. Jain, P. Kishore, P. Balamuralidhar, S.H. Damle, G. Viswanathan, Indian MST radar 1. System description and sample vector wind measurements in ST mode, Radio Sci. 30 (4) (1995) 1125–1138.
[5] M. Durga Rao, P. Kamaraj, J. Kamal Kumar, K. Jayaraj, K.M.V. Prasad, J. Raghavendra, T. Narayana Rao, A.K. Patra, The advanced Indian MST radar (AIR): system description and sample observations, Radio Sci. 55 (1) (2020).
[6] I.S. Rao, V.K. Anandan, Bispectral analysis of atmospheric radar signals, in: IEEE Aerospace and Electronic Systems Magazine, vol. 23, Aug. 2008, pp. 38–41, https://doi.org/10.1109/MAES.2008.4607898, 8.
[7] D. Uma Maheswara Rao, T.S. Reddy, G.R. Reddy, Atmospheric radar signal processing using principal component analysis, Digit. Signal Process. 32 (2014) 79–84.
[8] N.I. Eappen, T.S. Reddy, G.R. Reddy, Semiparametric algorithm for processing MST radar data, IEEE Trans. Geosci. Rem. Sens. 54 (5) (2015) 2713–2721.
[9] Raju, Analysis and evaluation of data-adaptive spectral estimation algorithms for processing MST radar data, Remote Sensing in Earth Systems Sciences 2 (4) (2019) 161–172.
[10] R. Bhatta, S.R.B. Biswas, Pillai, T.V.C. Sarma, Wind profile tracking in MST radar using viterbi data association, IEEE Trans. Geosci. Rem. Sens. 58 (10) (2020) 7074–7084.
[11] C. Raju, T. Sreenivasulu Reddy, MST radar signal processing using iterative adaptive approach, Geosci. 5 (T) (2018) 20. Lett.
[12] P.S. Babu, D.G. Sreenivasulu, Mesosphere stratosphere troposphere (MST) radar signal using discrete wavelet transform with overlapping group shrinkage, International Journal of Advanced Science and Technology 28 (9) (2019) 133–136.
[13] S. Thatiparthi, R. Gudheti, V. Sourirajan, MST radar signal processing using wavelet-based denoising, Geosci. Rem. Sens. Lett. IEEE 6 (4) (2009) 752–756.
[14] M. Crouse, R.D. Nowak, R.G. Baraniuk, Wavelet -based statistical signal processing using hidden Markov models, IEEE Trans. Signal Process. 46 (4) (1998).
[15] Marple Jr., S. Lawrence, William M. Carey, Digital Spectral Analysis with Applications, 1989.
[16] M.H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, 2009.
[17] D.L. Donoho, I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, J. Am. Stat. Assoc. 90 (432) (1995) 1200–1224.
[18] H.C. Huang, N. Cressie, Deterministic/stochastic wavelet decomposition for recovery of signal from noisy data, Technometrics 42 (3) (2000) 262–276.
[19] H.A. Chipman, E.D. Kolaczyk, R.E. McCulloch, Adaptive Bayesian wavelet shrinkage, J. Am. Stat. Assoc. 92 (440) (1997) 1413–1421.
[20] F. Abramovich, T. Sapatinas, B.W. Silverman, Wavelet thresholding via a Bayesian approach, J. Roy. Stat. Soc. B 60 (4) (1998) 725–749.
[21] M.R. Padhy, S. Vigneshwari, M.V. Ratnam, Application of empirical Bayes adaptive estimation technique for estimating winds from MST radar covering higher altitudes, Signal, Image and Video Processing (2023) 1–9.