Dictionary learning technique and penalized maximum likelihood for extending measurement range of a Rayleigh lidar

Varanasi Satya Sreekanth,[a,b,]* Karnam Raghunath,[a] and Deepak Mishra[b]

[a]National Atmospheric Research Laboratory, Department of Space, Gadanki,
Andhra Pradesh, India

[b]Indian Institute of Space Science and Technology, Thiruvananthapuram, Kerala, India

Abstract. Extending the temperature measurement range and sustaining the system performance are important requirements in Rayleigh lidars. While replacing hardware and increasing the power aperture product are ways of achieving these objectives, the data analysis approach is another means for achieving the same. A method to retrieve atmospheric temperatures using the penalized maximum likelihood method after denoising the backscattered signal using the dictionary learning technique is presented and compared with the conventional method. The proposed combination has the advantage of improving the measurement range and reducing the standard error (SE) in temperatures. The penalized maximum likelihood function is solved using the method of successive approximations, and the SE in temperature is calculated using Monte Carlo simulations. Observations from the Rayleigh lidar at the National Atmospheric Research Laboratory, India, are used for testing the approach. When compared with the conventional method, the SE in temperatures improved by 5K at 84 km, and the average height improvement was about 6 km. © 2020 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JRS.14.034529]

1 Introduction

The Rayleigh lidar system is an active remote sensing instrument used for atmospheric temperature measurement. The improvement in the signal-to-noise ratio (SNR) alone is not sufficient to retrieve temperature at higher altitudes. Therefore, a new method for temperature retrieval that uses prior is discussed. A comparative study on existing methods for temperature retrieval is discussed in the following paragraph.

The atmospheric temperature retrieval method proposed by Hauchecorne and Chanin (HC method)[1] is the most widely accepted method for temperature retrieval in the lidar community. The method uses the laws of physics and the downward integration for temperature retrieval. A seed pressure value is to be obtained from other instruments for initializing the downward integration process. An analytical expression for the uncertainty in the temperature retrieved, termed the standard error (SE), is given by Hauchecorne and Chanin in Ref. 1. They showed that this uncertainty decreases in proportion to the density at the top altitude to the density at each altitude below. In practice, when using the HC method, the top 10 to 15 km of the temperature profile is discarded due to uncertainty in seed pressure at the topmost altitude. This is a major disadvantage of the HC method. To overcome it, an inversion method was proposed by Sica and Jaya Khanna called grid search (GS) method[2] for atmospheric temperature retrieval. Using the mathematical inversion, the temperature values discarded are restored by providing the seed value from the lowest height rather than the highest altitude, where the variability of pressure is smaller in the mesosphere. The GS method extends the altitude range of existing Rayleigh-scatter lidars by 10 to 15 km, producing an equivalent of four times the power aperture product. All of the shortcomings in the above methods are overcome by the optimal estimation method

*Address all correspondence to Varanasi Satya Sreekanth, sreekanth@narl.gov.in

(OEM).[3] The OEM proposed by Sica and Haefele gives a complete uncertainty budget. The forward models presented in Ref. 3 completely characterize the measurements and allow for the retrieval of the temperature, dead time, and background counts. The method gives a full uncertainty budget, obtained on a perprofile basis, in addition to the statistical uncertainties. A procedure for atmospheric temperature retrieval using seed temperature values obtained from O_2 Airglow measurements is given by Taori et al.[4] Though this method reduces the SE by removing uncertainty in the seed parameter value, it requires colocated seed values, which puts a limitation on the method.

A detailed discussion on the upper altitude limit of the Rayleigh lidar is given in Refs. 5 and 6. Both the sodium[7,8] and potassium[9,10] temperature lidars utilize the Doppler broadening of metal atoms, so narrowband tunable lasers are required in the lidar transmitters to resolve the temperature-dependent Doppler-broadened atomic line shape. A temperature-sensitive ratio of the sodium fluorescence signals at two frequencies is used to derive temperature from the Doppler-broadening linewidth. The principle of deriving temperatures from a potassium lidar signal is very similar to the Na temperature lidar except the laser frequency was scanned through potassium D_1 line (instead of D_2) since the potassium D_2 line overlaps with a strong O_2 absorption line. The temperature dependence of the population ratio on different energy levels of atomic iron (Fe)[11–14] is used to derive temperatures in the mesosphere and the lower thermosphere (MLT) region. The temperatures are retrieved by taking the ratio of the fluorescence photon signals from two independent Fe lines. The measurement of temperature using a Fabry–Perot interferometer (FPI) through nocturnal airglow measurements is given in Refs. 17 and 18. The airglow instrument at NARL measures temperatures using the wavelengths of O_{1s}:558 nm, O_{1D}:630 nm (from 220 km), and OH:558 nm from 84 km. Nonlinear regression techniques when applied to sky exposures obtained using FPI give the temperature through the inversion of the resulting fringe pattern. A comparison of temperature retrieval from satellite instruments such as microwave limb sounder (MLS) and SABER-TIMED (Thermosphere Ionosphere Mesosphere Energetics Dynamics) instrument with ground-based lidars is given in Refs. 15 and 16.

The moving averages (MA) technique is the simplest denoising technique. Several filtering methods based on stationary and nonstationary signal processing such as Kalman filtering (KF),[19] discrete wavelet transform,[20–22] and empirical mode decomposition (EMD)[23,24] are employed by the users. KF, defined as a recursive algorithm, degrades the accuracy with abrupt changes in the retrieving parameter. A comparison of the SNR at different heights with the MA method, wavelet thresholding, and EMD for the Rayleigh lidar is given in Ref. 23. EMD method makes up for the limitations of wavelet thresholding EMD and KF. This method has good adaptability, thereby improving the SNR by a factor of 2.3. EMD has the disadvantage of a mode mixing problem that is solved by the variational mode decomposition (VMD).[25,26] The VMD adopts a new adaptive signal decomposition method that is superior to all variants of wavelet techniques with a different approach to mode decomposition. In the VMD technique, the selection of mode parameters is done through an optimization algorithm. The optimal parameter combination of mode number and quadratic penalty term improves the SNR by 29.3 dB. In recent years, the application of machine learning and neural networks for signal denoising has gained importance due to their adaptability to include prior information. Least squares support vector machine (LS-SVM)[27] and wavelet neural networks[28] are machine learning-based signal denoising techniques.

In this work, the photon count denoising is done through the dictionary learning technique. One of the key principles of dictionary learning is that the dictionary has to be inferred from the input data. Before this approach, the general practice was to use predefined dictionaries (such as Fourier or wavelet transforms). Using statistical information in the denoised photon count data, the penalized maximum likelihood ($KKTl_2$) method is used for retrieving atmospheric temperature. The SNR, temperature, and SE are compared with the conventional method, and the results are discussed. The detailed workflow path is shown in Fig. 1.

2 Dictionary Learning for Lidar Signal Denoising

The observed data matrix ($\mathbf{Y}$) is expressed as the sum of the linear combination of the product of the dictionary matrix ($\mathbf{D}$) with the sparse coefficient matrix ($\mathbf{X}$) and the noise matrix ($\mathbf{V}$). The term dictionary refers to a set of column vectors (basis vectors) called atoms.
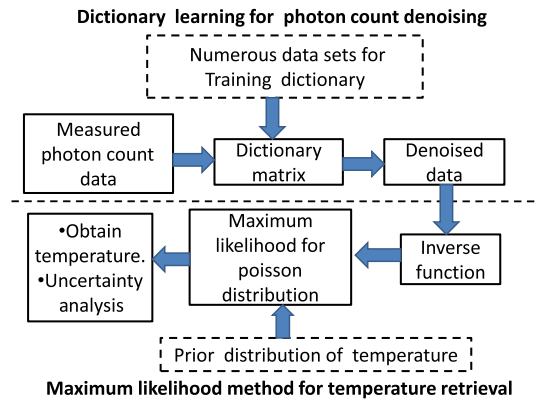
Fig. 1 Work flow diagram.

The basis vectors of the dictionary ($\mathbf{D}$) used to compute a sparse representation of the lidar data can be predefined using analytical functions that include transform techniques (sin and cos functions), Fourier series, and wavelets.[29–31] Using a predefined dictionary for the sparse representation of a lidar data set leads to a global approach, and the same dictionary is used regardless of the time or the offset of the data.

In this work, the dictionary matrix is learned from training the dictionary matrix, initialized with all zeros. Learning a dictionary matrix involves two phases: sparse coding and the dictionary update that is explained in detail in the below sections. Once the dictionary is learned, it can be used to denoise the photon count profiles through thresholding. The denoised photon count profiles obtained using the dictionary learning are inverted to retrieve temperature profiles using the $KKTl_2$ method.

Dictionary learning consists of two phases: sparse coding and the dictionary update. Sparse coding finds the sparse representation of the input data (also known as sparse coding) in the form of a linear combination of basic elements as well as those basic elements themselves. For the observed data $\mathbf{Y}$, there exists a dictionary $\mathbf{D}$ that represents observed data as the linear combination of a sparse matrix $\mathbf{X_i}$, with noise ($\mathbf{V}'$) as given in Eq. (1). In the case of the Rayleigh lidar, the matrices $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{D}$, and $\mathbf{V}'$ are analogous to the observed photon count matrix, sparse matrix, dictionary matrix, and random noise vector, respectively. Lower case bold letters represent the column vectors of the matrix. Dictionary learning with thresholding is used for random noise attenuation in this work:

$$[\mathbf{Y}]_{N\times m} = [\mathbf{D}]_{N\times K'}[\mathbf{X}]_{K'\times m} + [\mathbf{V}']_{N\times m}, \tag{1}$$

where N is the number of range bins, m is the number of profile integrations, and K' is an arbitrary size.

In this work, sparse coding is achieved through batch orthogonal matching pursuit (BOMP), and the dictionary update is achieved by K-singular value decomposition (K-SVD). K-SVD[32–35] is a generalization of the k-means clustering method, and it works iteratively by alternating between sparse coding of the observed data, based on the current dictionary, and dictionary update.

2.1 *Sparse Coding*

Sparse coding is a learning method that represents the observed data in the form of a linear combination of basic elements as well as those basic elements themselves. Sparsity corresponds to systems with few pairwise interactions. In numerical analysis and scientific computing, the sparse matrix or the sparse array is a matrix with most of the elements as zero. The sparse representation of $\mathbf{y}$ must satisfy $\|\mathbf{y} - \mathbf{Dx}\|_p \leq \delta$. The vector $\mathbf{x}$ contains the sparse coefficients of observed data. Here, $\|.\|_2$ represents the square root of the sum of squares of the individual elements. If $N < K'$ and $\mathbf{D}$ is the full rank matrix, an infinite number of solutions exist for obtaining the sparse coefficients.

The sparse coding phase is implemented using BOMP. BOMP has the advantage that it can work with large observed data effectively. The overcomplete basis has the advantage that it is

better capable of capturing structure and patterns in the input data. Each sparse coding step minimizes the representation error $\|\mathbf{Y} - \mathbf{DX}\|_2^2$. The BOMP algorithm is specifically optimized for sparse coding large sets of signals over the same dictionary and selects the atom with the highest correlation. The signal is orthogonally projected to the span of the selected atoms, the residual is recomputed, and the process is repeated. The procedure for implementation of the algorithm is given in Ref. 35.

2.2 Dictionary Update

2.2.1 K-singular value decomposition

There is a relation between sparse representation and clustering. The k-means clustering contributes to two steps in each iteration; given $d_k'^{K}$, it assigns training examples to their nearest neighbors and updates d_k'. Singular value decomposition (SVD) is the process of representing a given square matrix into the product of an upper triangular matrix (**U**), Eigenvalue matrix (Δ), and lower triangular matrix (**W**). The smaller magnitude of Eigenvalues and their corresponding vectors contribute to noise. A more detailed discussion about k-means clustering and SVD is given in Ref. 33.

The dictionary matrix **D** is initialized with zero as its elements or is initialized randomly by picking signals from the data signals. Each column of **D** is normalized by the square root of the sum of the squared elements of **D**. Minimization of Eq. (2) corresponds to searching the Dictionary for the sparse representation of observed data $\hat{\mathbf{Y}}$. The columns in **D** are fixed, and only one column in **D**, $\mathbf{d_{k'}}$ is updated at a time instant in the learning process. The product DX is decomposed into the sum of K' rank-1 matrices that approximates $\mathbf{E_{k'}}$. The matrix $\mathbf{E_{k'}}$ stands for the error for all N examples when the K'th atom is removed. The dictionary update process is explained in detail in Algorithm 1. Implementing Algorithm 1 one time forms a single iteration. The number of iterations will depend upon the error matrix $\mathbf{E_{k'}}$ and the tolerance value δ:

$$\min_{\mathbf{D,X}}\{\|\mathbf{Y} - \mathbf{DX}\|_2^2\} \quad \text{subject to} \quad \forall \ i, > \|\mathbf{x}_i\|_0 \leq T_0. \tag{2}$$

where i is an arbitrary variable, and T_0 is logic-1(True).

Algorithm 1 Sparse coding and dictionary update.

1: **Procedure** DICTIONARY LEARNING

2: Initialization: Choose the photon count data with SNR $\geq$ 3.

3: **Sparse Coding:**

4:     Implement Eq. (2)

5:     Decompose **DX** into the sum of K' matrices with rank 1.

6: **Dictionary Update:**

7: For each column k' in **D**, update using $\omega_{k'}$ and $\mathbf{x}_{\mathbf{k'}}^{T'}$. $\omega_{k'}$ defines a group of indices pointing $\mathbf{y}_i$. $\omega_{k'} = i | 1 < i < k'$, $X_{k'}^{T'}(i) \neq 0$, omega_k' is an arbitrary variable, and T' is transpose of the matrix.

8: Calculate representation error matrix $\mathbf{E_{K'}}$. The expression for the representation error is given by $\mathbf{E_{k'}} = \mathbf{Y} - \sum_{j \neq k'} \mathbf{d_j X_{T'}^j}$, where **j** is an arbitrary variable. Restrict $\mathbf{E_{k'}}$ by choosing only columns corresponding to $\omega_{k'}$ and $\mathbf{E_{k'}^R}$. Get $\mathbf{E_{k'}^R}$ by choosing a column in $\mathbf{E_{k}'}$ (nearest neighbor).

9: Calculate SVD for $\mathbf{E_{k'}^R}$. Arrange eigenvalues of $\mathbf{E_{k'}^R}$ in descending order and replace eigenvectors corresponding to the smallest value of eigenvalues by zeros in $\mathbf{E_{k'}^R} = \mathbf{U} \Delta \mathbf{W^{T'}}$, where bold letter **U** is the upper triangular matrix, Δ is the Eigen matrix, and **W** is the lower triangular matrix.

10: Update $\mathbf{d_k'}$ with the first column of **U** and $\mathbf{x_{k'}^R}$ to be the first column of **W**. Eliminate atoms that are close to each other.

11: Eliminate atoms that are used by less than b (arbitrary value) training examples.

2.3 *Thresholding*

There are two major sources of noise in the Rayleigh lidar system, namely random noise and background noise.

- Attenuating random noise [$v' \sim \mathcal{N}(0, \sigma'^2)$] using sparsity is achieved by computing the sparse approximation in the dictionary domain, where σ' is the variance of the random noise. The noise cannot be represented by the sparse approximation, and it is attenuated.
- Sky background noise is separated from the observed photon counts by removing the average of all photon counts from a certain range bin to the maximum number of range bins.
- Denoising observed data: The denoised data are recovered by applying thresholding in the learned dictionary domain. The observed data are the sum of true data and noise:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{v}', \tag{3}$$

where $\mathbf{y}$ denotes the observed data and $\hat{\mathbf{y}}$ denotes the true data. The σ' value is selected by choosing trails that give a good SNR at higher altitudes. The true data have a sparse representation in dictionary basis. So $\overline{\mathbf{y}} = \mathbf{D}^{\mathbf{T}}\mathbf{y} = \mathbf{D}^{\mathbf{T}}\hat{\mathbf{y}} + \mathbf{D}^{\mathbf{T}}\mathbf{v}' = p' + z'$, where $\overline{\mathbf{y}}$, p', and z' are arbitrary variables. Most elements in p' are 0 or close to 0 and $z' \sim \mathcal{N}(0, \sigma'^2 I)$. Since $\mathbf{D}$ is orthogonal, the estimation problem amounts to the recovery of a signal in i.i.d. Gaussian noise. Select the prior as $p'' \sim a\mathcal{N}(0, \sigma_1'^2) + (1 - a)\mathcal{N}(0, \sigma_2'^2)$, where p'' is Prior for denoising, and $\mathbf{I}$ is the identity matrix, a is the arbitrary value, $\sigma_1'^2$ is the variance of "significant" coefficients, and $\sigma_2'^2$ is the variance of "insignificant" coefficients. Then, $\tilde{p}' = E'(p''/y) = \tau(y)y$, where $E'(.)$ is the expectation of function and $\tau(y)$ is the shrinkage factor, which depends on the prior variances $\sigma_1'^2$ and $\sigma_2'^2$. The effect of the shrinkage factor is that small coefficients are set early to 0 and large coefficients are unaltered. Small coefficients are mostly noises, and large coefficients contain actual signal. Apply the inverse transform to obtain the denoised data $\hat{\mathbf{y}} = \mathbf{D}\tilde{\mathbf{p}}'$.

2.4 *Test Metric for the SNR Improvement*

The original photon count profiles and the denoised photon count profiles with the dictionary learning technique follow the Poisson distribution in clear sky conditions. The test metric for improvement in the SNR is symmetric Kullback–Leibler (SKL) divergence [Eq. (5)]. The mathematical expression for SKL divergence between signal and noise is given as

$$\mathrm{KL}(P'||Q') = \sum_i P'(i) \ln\left[\frac{P'(i)}{Q'(i)}\right], \tag{4}$$

$$\mathrm{SKL}(P'||Q') = \frac{\mathrm{KL}(P'||Q') + \mathrm{KL}(Q'||P')}{2}, \tag{5}$$

where KL is the Kullback–Leibler divergence and P' and Q' indicate the distribution of signal counts and noise counts, respectively, for a technique. A higher value of SKL indicates that signal counts are increasing distinctly from noise. SKL divergence is zero when signal and noise counts are equal ($P' = Q'$), and it peaks near the source for all denoising techniques. It falls at a slower rate for the top-performing technique.

3 Temperature Retrieval Using Penalized Maximum Likelihood (KKTl$_2$) Method

The expression for lidar back scattered signal[37] is of the form:

$$P^*(z) = d'(z) \exp\left[-\int_0^z \alpha(x)\mathrm{d}x\right], \tag{6}$$

where $P^*(z)$ is the power of the backscattered signal, $d'(z)$ is the system-dependent function, $\alpha(x)$ is the parameter related to the temperature, and the exponential term gives the

transmission coefficient. The forward model [Eq. (9)] is used for inverting the temperature profile:

$$Y = F(X) + V', \tag{7}$$

$$\rho(z_i) = C[n(z_i) - B](z_0 - z_i)^2, \tag{8}$$

$$n(z_i) = \frac{P(z_i)M_{\text{air}}}{CRz_i^2 T(z_i)} \exp\left[\int_{z_0}^{z_i} \frac{-M_{\text{air}}g(z)}{RT(z)} \, dz\right], \tag{9}$$

where $P(z_i)$ and $T(z_i)$ are pressure and temperature at altitude z_i, $\rho(z_i)$ is the atmospheric density at the i'th layer, and $F: \Re \to \Re$ is a nonlinear operator. $\Re$ indicates the set of real numbers. Equation (9) is the forward model[37] that gives the relation between observed photon counts $[n(z)]$ and the temperature parameter, where C is the lidar system constant, B is the background photon count, z_0 is the altitude of the top layer, $g(z)$ is the acceleration due to gravity, R is the universal gas constant, M_{air} is the molar mass of dry air, and z is the height above the Earth's surface. The constant C or the normalization constant is a combination of all of the constants occurring in the simplified lidar equation. It depends on the laser power, the optical and quantum efficiencies of the system, and the two-way transmission coefficient.

3.1 Maximum Likelihood for Poisson Data

It is convenient to express [Eq. (9)] in its discrete form as

$$P^* = d' \odot \exp\{-L\alpha^*\}, \tag{10}$$

where $\odot$ is the elementwise multiplication also known as the Hadamard Product, P^* is the observed vector such that P_i^*, $i = 1, 2, 3, \ldots N$, $d'(z_i)$ is the unknown system function vector $i = 1, 2, 3, \ldots, N$, $\alpha_j^* = \alpha(z_j)$, $j = 1, 2, \ldots, m$, L is the operator discretizing the integral. From here onward, it is assumed that the noise is Poisson distributed. The noise at each range bin is independent of other bins. We aim to recover the temperature values from the backscattered signal $P^*(z)$. Here, α is a variable that depends on temperature. $\exp(-L\alpha)$ constitutes the atmospheric transmission coefficient.

3.2 Iterative Algorithm for Temperature Retrieval

Maximizing Eq. (10) that solve for α

$$S(\alpha) := J(\alpha). \tag{11}$$

The maximum of log-likelihood exists and is unique provided that $N \geq m$. $L^{T'}$ is thought of as a cumulative weighted sum from the top height. The first-order KKT (Karush–Khun–Tucker) conditions are necessary conditions for the optimality under given constraints. The sign of the derivate of $S(\alpha)$ gives information about the existence of minima or maxima. Equation (11) is solved using the iterative algorithm.

The maximum likelihood function for the photon count and its detailed solution is given in Ref. 38. The final solution to Eq. (11) is given as

$$\alpha^{(n+1)} := \frac{L^{T'}[d' \odot e^{-L\alpha^{(n)}}]}{(L^{T'}P^*)} \odot \alpha^{(n)}. \tag{12}$$

By comparing Eqs. (9) and (10), we have $P^*(z_i) = n(z_i)$, L is an operator discretizing the integral, $d'(z_i) = \frac{P_{(z_i)}M_{\text{air}}}{CRz_i^2 T(z_i)}$, and $\alpha(z) = \frac{M_{\text{air}}g(z)}{RT(z)}$. The optimization problem is the unconstrained, and the initial value of α is required for iteratively deriving the value of temperature at all range bins. The initial procedure is as follows:
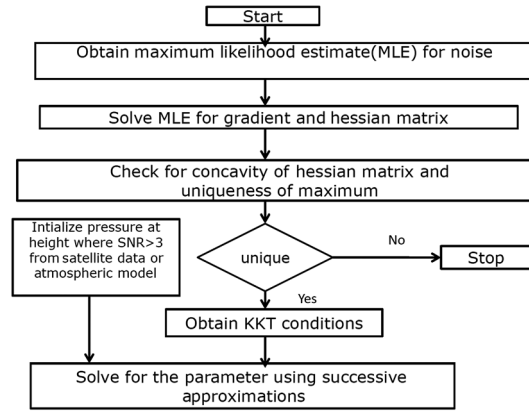
Fig. 2 $KKTl_2$ algorithm.

1. Identify the value of N such that $P^*(z_N) \neq 0$.
2. Calculate the value of α components by performing the line-search method after obtaining the initial value of temperature.

3.3 Penalized Maximum Likelihood ($KKTl_2$)

The method is to have an l^2 penalty term weighted by the regularization parameter γ. The expression for likelihood function with penalty term is given as

$$\hat{S}(\alpha) = J(\alpha) - \gamma\|\alpha\|_2^2. \tag{13}$$

Equation (13) is solved using the procedure followed for KKT-maximum likelihood. The solution to it is given in Ref. 38 as

$$\alpha^{(n+1)} := \frac{L^{T'}[d \odot e^{-L\alpha^{(n)}}]}{[L^{T'}P^* + 2\gamma\alpha^{(n)}]} \odot \alpha^{(n)}. \tag{14}$$

The term $2\gamma\alpha^{(n)}$ is the weighting term that reduces the error at the higher altitudes. The regularization parameter (γ) is considered the inverse of the variance of the prior distribution. The prior profile for initialization of the algorithm is obtained from atmospheric model MSISE-00, Rayleigh lidar, sodium, and potassium resonance lidar, Fe–Boltzmann lidar, and satellite instruments such as MLS and SABER. So the regularization parameter does not depend on the SNR. The sequence of steps followed in the implementation of the $KKTl_2$ method is shown in Fig. 2.

4 Numerical Experiments

The Rayleigh lidar system at NARL (13.5°N, 79.2°E) has been in operation since 1998. The Rayleigh lidar at the National Atmospheric Research Laboratory (NARL), Gadanki, India, operates with a wavelength of 532 nm at 600 mJ of energy per pulse with a pulse repetition rate (PRR) at 50 Hz. The receiving system consists of a Newtonian type telescope with a diameter of 750 mm and field of view of 1 mrad. A photomultiplier tube (PMT) is connected at the focal point of the telescope and converts optical signals to electric signals. The output of the PMT is fed to pulse discriminators, and a multichannel scaler is used for acquisition and photon counting. The Rayleigh lidar system consists of two channels, namely a high gain channel and a low gain channel for extending the dynamic range of the system. The low gain channel is meant for collecting backscattered signals from lower altitudes (<50 km) and a high gain channel for higher heights (>50 km). The present interest is extending the measuring range of the Rayleigh lidar, so only the high gain channel is considered in further discussions for signal analysis. The two channels have two PMTs with identical specifications, and their gains are in the ratio

of 9:1 for high gain and low gain channels, respectively. The temperature profiles are retrieved for each channel and merged in the vertical height range between 45 km and 55 km using the expression given in Ref. 36. The data used for analysis are collected for 1 h.

4.1 *Application to Rayleigh Lidar*

The SNR in the case of Poisson error in the simplest form is defined as the ratio of received photon counts (S) due to scatters and the square root of the sum of the source counts (S) and background count (B). The background count is calculated by taking an average of all photon counts in range bins between 150 and 300 km:

$$\text{SNR} = \frac{S}{\sqrt{S+B}} \approx \sqrt{S}, \quad S \gg B. \tag{15}$$

The system, which has shown significant values of the SNR up to 90 km in the initial years, has gone down to 78 km over the course of time. A standard procedure for increasing the SNR is by coadding the photon count profiles in space and time in clear sky conditions or increasing the laser energy and receiver aperture size. Adding data in time series improves the SNR by the square root of the sample size. However, certain features will be missing due to more time integrations. Data in January 1999 have a high SNR as the system was recently installed, and hence they are used for training the dictionary. Each photon count profile is collected with a range resolution of 0.3 km and a time integration of 4 min. The dictionary was learned with the data set of size 200×450, of which 450 is the number of profiles used for training and 200 is the number of range bins.

Photon count data with a high value of SNR are used for training the dictionary. The data used for learning are divided into small patches to increase computational efficiency. The size of the data matrix used for training must be an integer multiple of the patch size to overcome patch overlap problems. In this work, the dictionary is trained using data with a high SNR acquired during 1998 and 1999. The patch size is of dimension 16×16, and K-SVD is iterated 80 times. Here, we chose 80 iterations in which the representation error tolerance reaches $\delta = 0.001$ to get the optimal set of the dictionary atoms. Noise variances of $\sigma' = 3$, $\sigma_1' = 4$, $\sigma_2' = 8$, and $a = 0.5$ are taken in training the dictionary. SKL metric and uncertainty analysis are performed for data collected on January 5, 2011.

The dictionary matrix, the sparse coefficient matrix, the noise matrix, and the denoised data matrix computed for the data collected on January 5, 2011, are shown in Fig. 3. Values of $N = 200$ and $m = 15$ are used while denoising the observed photon count profiles. The size K' varies depending upon the number of profiles used for learning the dictionary. The value of $k' = 324$ is obtained after training the dictionary with data of size 200×450. In Fig. 3(b), the dictionary coefficients are at the maximum value at lower heights and between the 175th and 190th range bins in all of the profiles. The reason for the maximum at lower range bins is due to the high SNR and that for the maximum coefficients between the 175th and 190th bins is due to improvement of the SNR at higher altitudes. The SNR for the DL technique is calculated as

$$\text{SNR}_{\text{DL}} = \sqrt{\left(\frac{\hat{\mathbf{y}}}{\mathbf{y} - \hat{\mathbf{y}}}\right)}. \tag{16}$$

The improvement in the SNR using the dictionary learning technique Eq. (16) is compared with the conventional technique [Eq. (15)] shown in Fig. 4. The specific days chosen are those in which the decrease in the measurement range is observed. The SNR at a given height has been decreasing over the years as observed from Tables 1 and 2. For a given threshold (SNR ≥ 3), the dictionary learning technique shows the same value of the SNR at the higher heights as the conventional technique. This height varies from a minimum of 4 km to the maximum of 6 km, as observed from Table 2.

To show the improvement in the SNR, comparison of the SKL metric between the DL technique and conventional technique is shown in Fig. 5. From Fig. 5, it is observed that the SNR of the dictionary learning technique falls at a slower rate than the conventional technique. The SKL
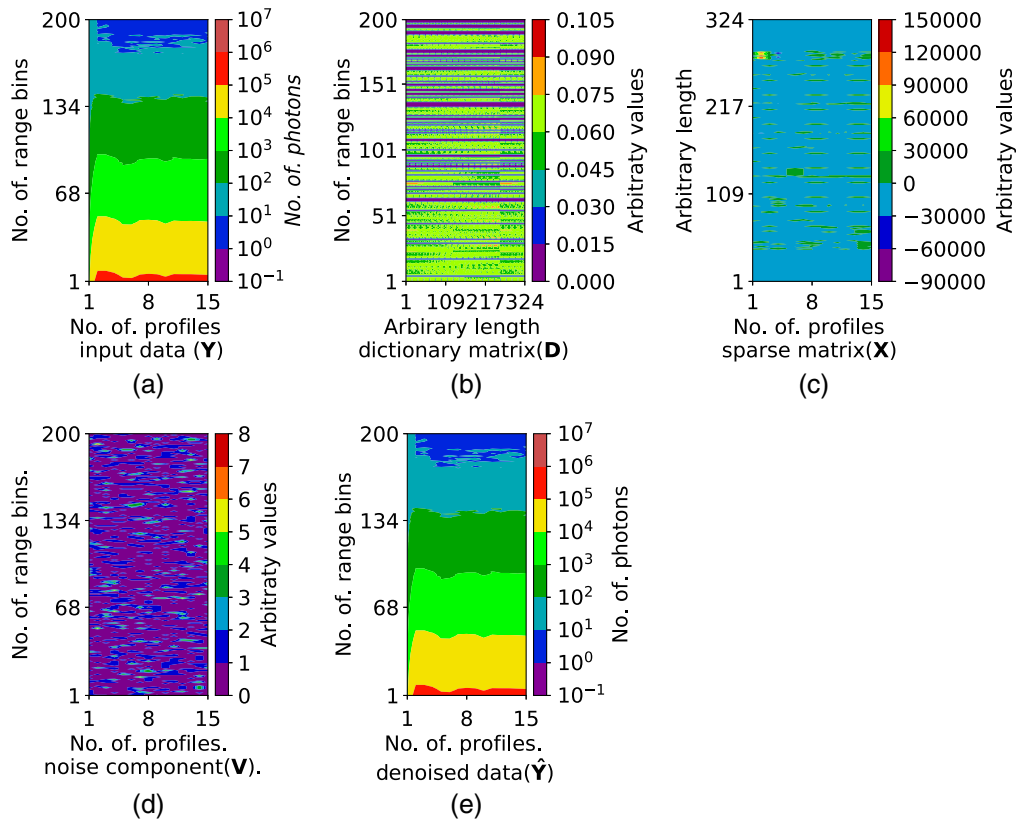
Fig. 3 Final matrices after the dictionary learning for denoising. (a) Input photon count matrix collected on January 5, 2011, with size of 200×15, (b) dictionary matrix after learning with size of 200×324, (c) sparse matrix with size of 324×15, (d) noise component matrix with size of 200×15, and (e) denoised data matrix with size of 200×15.

metric score for both techniques is zero after 92 km, where the signal and noise counts are indistinguishable in clear sky conditions.

There is no analytical expression for the prior term $(2\gamma\alpha)$ estimated as a function of altitude. The contribution of the prior is analyzed using simulation.

- To the observed photon count profile (true profile), a structure of the photon count profile is added in higher range bins, and the resultant profile is called the estimated profile. Adjusted temperatures obtained using the observed photon profile and the estimated photon count profile using the $KKTl_2$ are exact.
- The initialization parameters are selected in such a way that the discrepancy between exact and the estimated temperature profiles is minimum.
- To assess the sensitivity, 1000 Poisson realizations with the same mean of the estimated photon profile are generated, and the estimated temperature profiles for all 1000 realizations are computed using the $KKTl_2$ algorithm. Their variance is minimum in comparison with the exact profile. This indicates that the $KKTl_2$ manages to recover the presence and intensity of structure at higher altitudes.

4.2 Error Estimation in $KKTl_2$ Method for Temperature Retrieval

The $KKTl_2$ method is a numerical method, and there is no analytical expression for the SE. The error in the temperature retrieved is analyzed using the Monte Carlo method.[38,39] The contribution of the errors due to the seed pressure[40] is more than all other sources errors that are due to the instrument, counting, and geophysical variabilities. The seed pressure value is taken from the SABER TIMED instrument. A 10% change in seed pressure is used to analyze seasonal
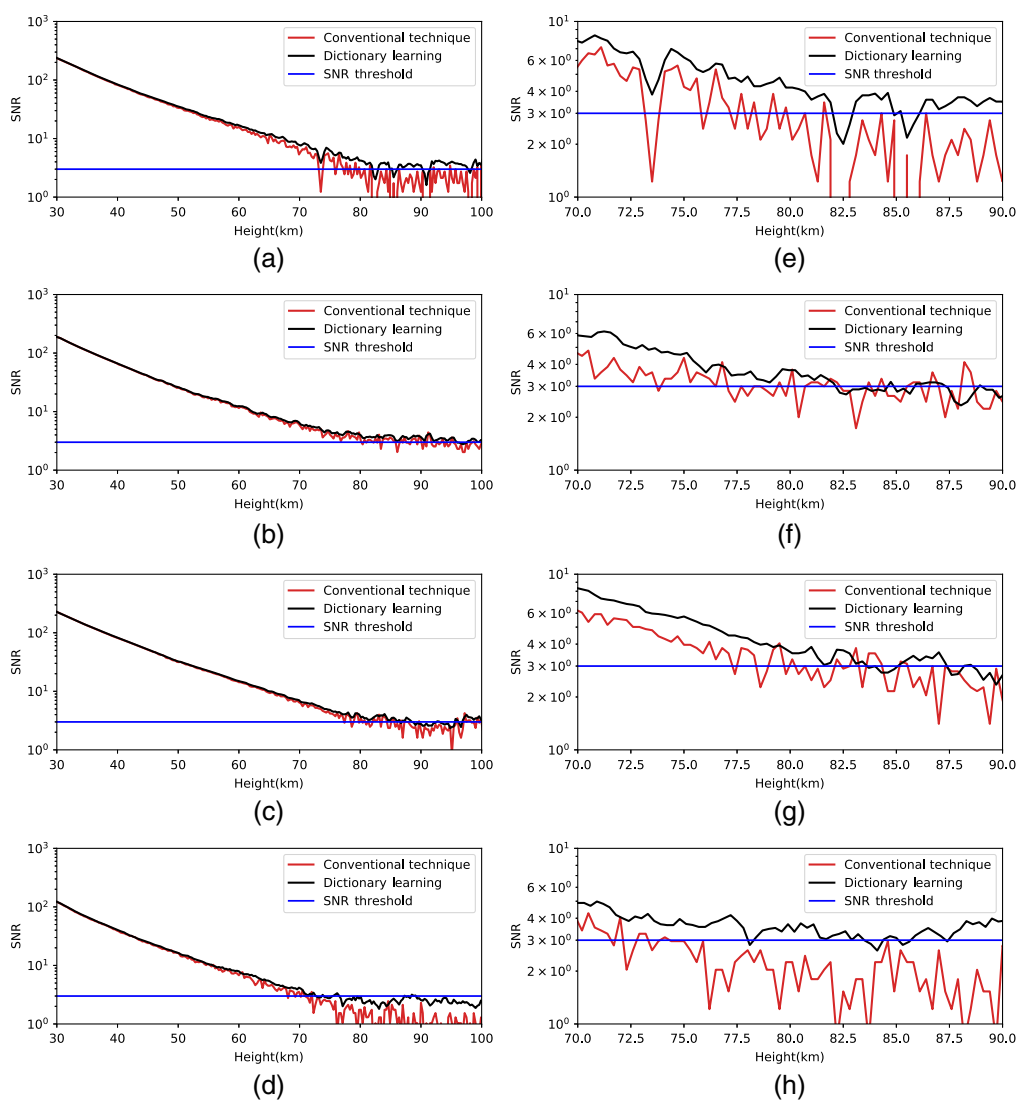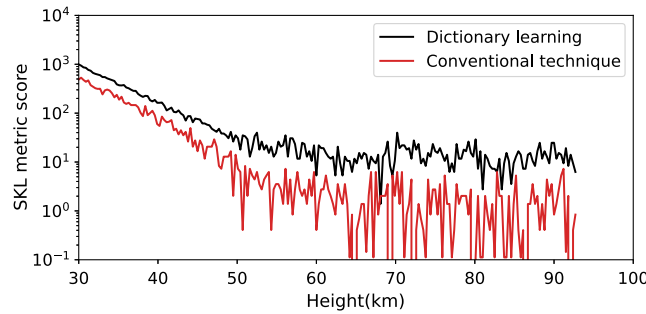
Fig. 4 Comparison of SNR using conventional technique and the dictionary learning for 1 h integration time with range resolution of 300 m on (a), (e) December 28, 1998, (b), (f) January 11, 1999, (c), (g) January 16, 2008, and (d), (h) January 5, 2011.

Table 1 Comparison of SNR.

Date		Height (km)				
		30	45	60	75	85
December 28, 1998 (Energy = 530 mJ, PRR = 20 Hz)	SNR	198	45	12	4	2
	SNR_{DL}	225	53	14	5	4
January 11, 1999 (Energy = 290 mJ, PRR = 20 Hz)	SNR	135	34	8	4	1
	SNR_{DL}	166	37	11	5	2
January 16, 2008 (Energy = 310 mJ, PRR = 50 Hz)	SNR	183	48	15	4	1
	SNR_{DL}	229	53	16	6	3
January 5, 2011 (Energy = 580 mJ, PRR = 50 Hz)	SNR	102	22	5	3	1
	SNR_{DL}	105	29	6	4	2

Table 2 Comparison of heights where SNR is 3.

Date	December 28, 1998	January 11, 1999	January 16, 2008	January 5, 2011
Height (km)	81	78	78	75
Height$_{DL}$ (km)	87	82	84	78

Fig. 5 Comparison of SKL metric between the dictionary learning and the conventional technique for 1 h integration time with range resolution of 300 m on January 5, 2011.

variations as mentioned in Ref. 40. Synthetic photon count profiles are generated with a random number generator that produces Poisson distributed random numbers with the mean as the actual photon count data at each altitude range, with a sample size of 150. To model the error due to seed pressure (P_0), seed pressure is modeled as a random variable that follows the normal distribution $\mathcal{N}(\mu, \sigma^2)$ with a mean (μ) as $P_{0_{true}}$ and standard deviation (σ) such that the $3\sigma = 10\% P_{0_{true}}$, which means 99.73% of changes will lie in the confidence interval $P_{0_{true}} + 10\% P_{0_{true}}$. A sample of 150 temperatures is obtained by randomly choosing the seed pressure value from a normal distribution with mean $P_{0_{true}}$ and standard deviation (σ). The standard deviation of all temperatures at each range bin is the error due to seed pressure in temperature using the $KKTl_2$ method. The algorithm is initialized with a pressure of 2.093×10^{-3} hPa at 95 km and is taken from the satellite instrument SABER-TIMED. The top 5 km of the temperature profile, where the effect of seed pressure on uncertainty in temperature is high, is discarded. The sounding of the atmosphere using the broadband emission radiometry (SABER) instrument is one of four instruments on NASA's TIMED (Thermosphere Ionosphere Mesosphere Energetics Dynamics) satellite with a vertical resolution of 0.4 km. The temperature profile is selected in such a way that the satellite has an overpass near the Gadanki location within the area of 100 km. Seed pressure is varied by picking values from a Gaussian random number generator with mean of $P_{0_{true}}$ and standard deviation of $3.33\% P_{0_{true}}$. The uncertainties in the temperature retrieved due to seed pressure are shown in Fig. 6 for the observed data on January 5, 2011. The uncertainty variations taken for this study are valid as temperatures derived from actual data and synthetic data are identical. The uncertainty is less by 5K between 70 to 85 km when compared with the GS and OEM methods. Further studies are required to have the complete uncertainty estimate of the vertical height and the temperature in tune with international lidar network station standards.[41–43]

The expressions for temperature [Eq. (17)] and SE [Eq. (18)] are given in Ref. 1 for the HC method. A seed pressure value from an atmospheric model is used for the initialization of the method for temperature retrieval:

$$T(z) = \frac{M_{air}g(z)dz}{R \log(1 + X')}; \quad X' = \frac{\rho(z)g(z)dz}{P(z + \Delta z/2)}, \tag{17}$$

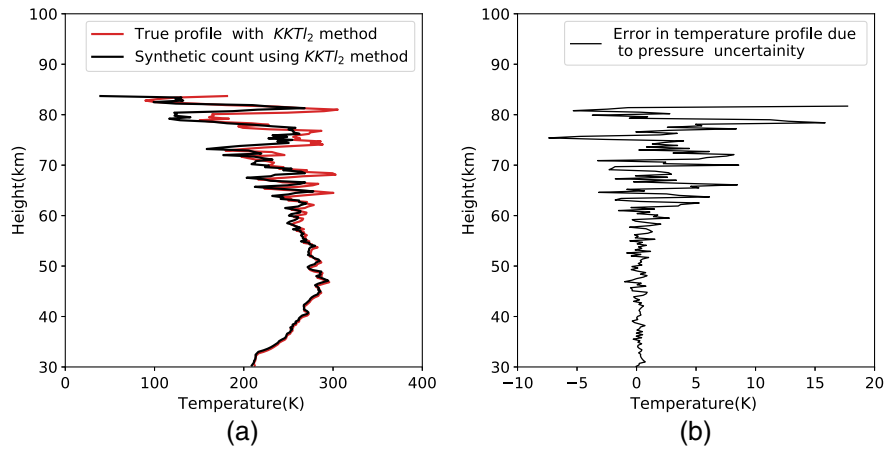$$SE = \frac{dX'}{(1 + X')\log(1 + X')}. \tag{18}$$

Fig. 6 A sample of 150 temperatures is obtained by randomly choosing the seed pressure value (P_0) from a Gaussian distribution with a mean of $P_{0_{true}}$ and standard deviation of $\sigma = 3.33\%P_{0_{true}}$, where $P_0 = 2.093 \times 10^{-3}$ hPa at 95 km: (a) synthetic count temperature profile and (b) uncertainty in temperature due to seed pressure uncertainty for January 5, 2011.

5 Discussions

For each data set, the temperature and the SE using the HC method and the dictionary learning technique in combination with the $KKTl_2$ method are compared in Fig. 7. The increase in the measurement range is noted. The average of all cases is taken to arrive at the overall increase in the measurement range of the instrument. The algorithm is tested with different seed pressure values.
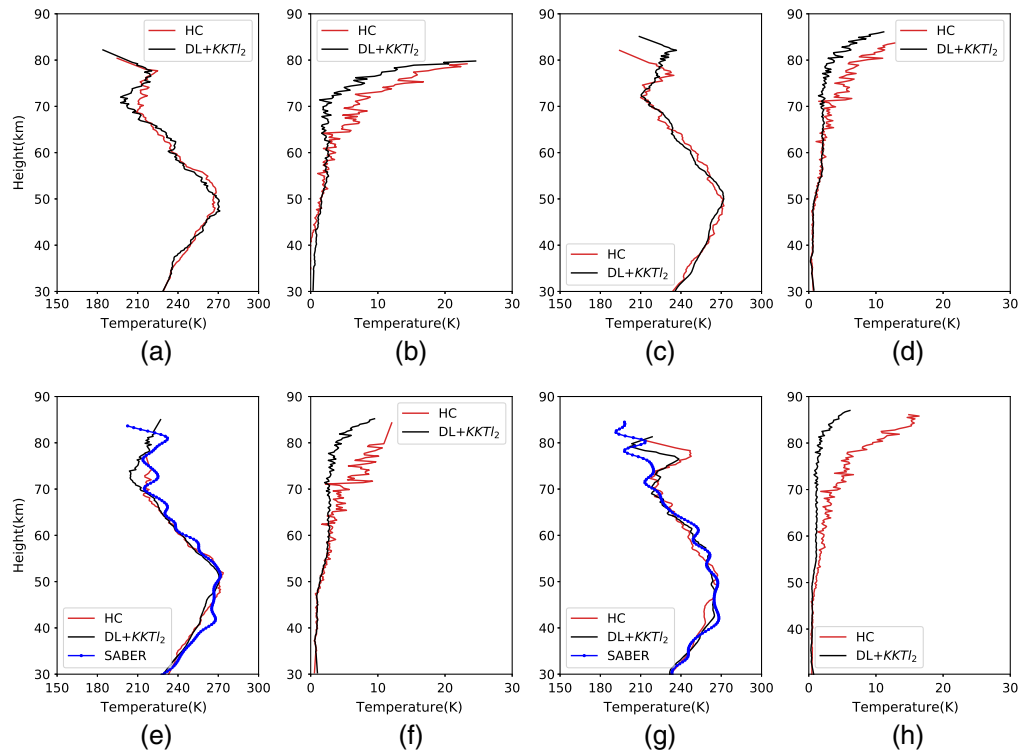

Fig. 7 Comparison of (a), (c), (e), (g) temperature and (b), (d), (f), (h) SE using HC, the dictionary learning technique for denoising, and the $KKTl_2$ method for December 28, 1998, January 11, 1999, January 16, 2008, and January 5, 2011, initialized with seed pressure of 2.093×10^{-3} hPa at 95 km.

The temperature profiles retrieved using the $KKTl_2$ method are compared with the temperatures retrieved using the HC method and the SABER. The results are shown in Figs. 7(a), 7(c), 7(e), and 7(g) and that of error in Figs. 7(b), 7(d), 7(f), and 7(h). Improvement in the measurement range is seen with respect to the SE. For a specific value of the SE, error in the HC method occurs at a height lower than that of the height obtained using the combination of the dictionary learning technique and the penalized maximum likelihood method. The error difference with the two methods is <6 K up to 48 km on all days as observed from Fig. 7. There is an improvement in the measurement range by 6 km at higher altitudes where the SNR is less. An SE of 8K is observed at 78 and 84 km with the HC method and the $KKTl_2$ method, respectively. On applying the $KKTl_2$ method to large data sets on different years, a minimum of 4 km and a maximum of 7 km improvement in measurement range is seen, resulting in an average height improvement of about 6 km. It is observed from Figs. 7(e) and 7(g) that the temperature retrieved by the $KKTl_2$ method is warmer than the temperature measured with SABER in the stratopause region by 5K.

An ensemble of temperature profiles retrieved using the $KKTl_2$ method and the temperature differences between the $KKTl_2$ method and the HC method are shown in Figs. 8(a) and 8(b), respectively. The ensemble of temperature profiles is for 1 h integration time observed data in clear sky conditions from March 1998 to March 2020, 2 days in a month with a range resolution of 0.3 km. From Fig. 8(b), it is observed that the temperatures retrieved with $KKTl_2$ are colder than the temperatures retrieved with the HC method, and the difference varies from 0K to 3K for most of the profiles and a maximum of 12K in the height range of 70 to 85 km. In the height range of 55 to 70 km, the temperatures retrieved with the $KKTl_2$ method are warmer than the temperatures retrieved with the HC method. The temperatures retrieved with $KKTl_2$ are colder than the temperatures retrieved with the HC method at the stratopause. Ideally, the difference between T_{KKTl_2} and T_{HC} is zero in the stratosphere where SNR $\gg$ 1. However, the temperature difference $(T_{KKTl_2} - T_{HC})$ is <3 K in the altitude range of 30 to 40 km and <6 K up to 70 km. The difference in the mesosphere could be the change of photon counting distribution and the height of seed pressure initialization. Investigations are needed to understand the sources of the discrepancies and to improve the methods to decrease the difference to <0.5 K in the stratosphere.
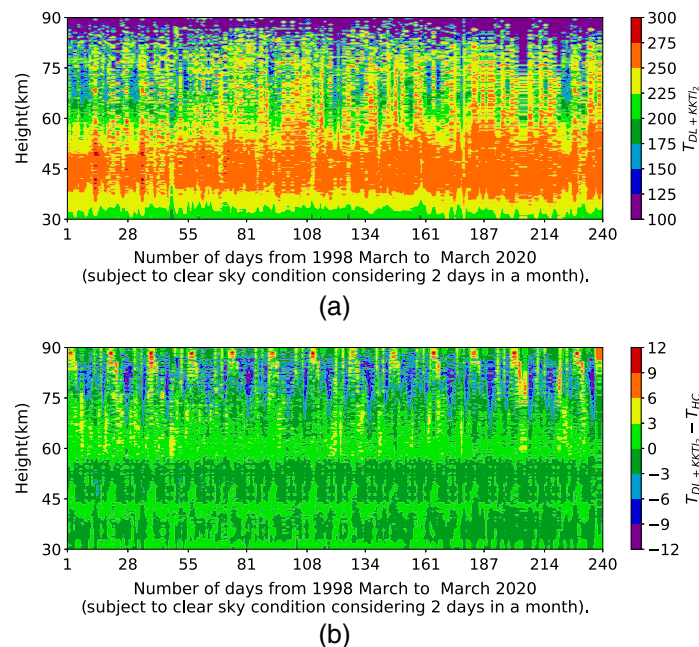

(a)

(b)

Fig. 8 Ensemble of temperature profiles and the temperature difference between the $KKTl_2$ and the HC methods.

6 Conclusion

The reduction in the measurement range of the Rayleigh lidar system over time is addressed using signal processing techniques, and useful results have been obtained. The main highlights of this work are:

- The dictionary learning (with the sparse coding stage using BOMP and the updating of the dictionary with K-SVD) technique is used for Rayleigh lidar signal denoising.
- An iterative method for atmospheric temperature retrieval is presented by solving a penalized maximum likelihood method to reduce the SE.
- The uncertainties in the retrieved temperatures are calculated using the Monte Carlo simulations.
- Both methods in combination improved the measurement range of the Rayleigh lidar system by 6 km.
- There is a reduction in the SE when photon count data are denoised using the dictionary learning technique in combination with the $KKTl_2$ method.

7 Appendix A: Machine Learning Terms

The following are basic terms that are used in machine learning terminology.

- Atoms: Any data are represented in the form of a linear combination of basic elements as well as those basic elements themselves. These elements are called atoms, and they compose the dictionary.
- Span: In linear algebra, the linear span of a set $\mathbf{S}$ of vectors in a vector space is the smallest linear subspace that contains the set.
- Residual:The residual is defined as the difference between the observed value of the dependent variable ($\mathbf{y}$) and the predicted value ($\hat{\mathbf{y}}$).
- Sparsity: In numerical analysis and scientific computing, a sparse matrix or sparse array is a matrix in which most of the elements are zero. By contrast, if most of the elements are nonzero, then the matrix is considered dense. The number of zero-valued elements divided by the total number of elements is called the sparsity of the matrix (that is equal to 1 minus the density of the matrix).
- Convolutional sparse coding: If the product in Eq. (2) is replaced by a convolutional operator, then the process of finding sparse codes is called convolutional sparse coding.
- Rank of a matrix: The rank of a matrix is defined as (a) the maximum number of linearly independent column vectors in the matrix or (b) the maximum number of linearly independent row vectors in the matrix. Both definitions are equivalent. For an $r \times c$ matrix, if r is less than c, then the maximum rank of the matrix is r.
- Matching pursuit: Matching pursuit (MP) is the sparse approximation algorithm that finds the "best matching" projections of multidimensional data onto the span of an overcomplete (i.e., redundant) dictionary D. The basic idea is to approximately represent a signal y from Hilbert space H as a weighted sum of finitely many functions g_{γ_n} (called atoms) taken from D. An approximation with N atoms has the form

$$y(t) \approx \hat{y}_N(t) := \sum_{n=1}^{N} a_n g_{\gamma_n}(t) y(t) \approx \hat{y}_N(t) := \sum_{n=1}^{N} a_n g_{\gamma_n}(t).$$

Normally not all of the atoms in $\mathbf{D}$ are used for signal representation. MP algorithms choose the atoms one at a time to reduce the approximation error. This is achieved by finding the atom that has the highest inner product with the signal ($\mathbf{y}$) (assuming the atoms are normalized), subtracting from the signal an approximation that uses only that one atom, and repeating the process until the signal is satisfactorily decomposed

- Support vectors machine (SVM): SVM is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in

n-dimensional space (where n is the number of features you have) with the value of each feature being the value of a particular coordinate.

- Least square support vector machine (LS-SVM): LS-SVM is the least-squares versions of SVM and are a set of related supervised learning methods that analyze data and recognize patterns for classification and regression analysis.
- Eigenvalue: For a given arbitrary square matrix $\mathbf{A}$ and identity matrix $\mathbf{I}$,

$$|\mathbf{A} - \lambda\mathbf{I}| = 0, \tag{19}$$

is called a characteristic equation. The solution to a characteristic equation gives eigenvalues.

The following is the expression for the error matrix ($\mathbf{E_k}$) obtained from the sparse expression:

$$\|\mathbf{Y} - \mathbf{DX}\|_F^2 = \left\|\mathbf{Y} - \sum_{j=1}^{K'}\mathbf{d}_j\mathbf{x}_T^j\right\|_F^2 = \left\|\left(\mathbf{Y} - \sum_{j\neq k'}\mathbf{d}_j\mathbf{x}_T^j\right) - \mathbf{d}_{k'}\mathbf{x}_T^{k'}\right\|_F^2 = \|\mathbf{E}_{k'} - \mathbf{d}_{k'}\mathbf{x}_T^{k'}\|_F^2. \tag{20}$$

Acknowledgments

References

1. A. Hauchecorne and M. L. Chanin, "Density and temperature profiles obtained by lidar between 35 and 70 km," *Geophys. Res. Lett.* **7**(8), 565–568 (1980).
2. J. Khanna et al., "New technique for retrieval of atmospheric temperature profiles from Rayleigh-scatter lidar measurements using non-linear inversion," *Appl. Opt.* **51**(33), 7945–7951 (2012).
3. R. J. Sica and A. Hefeli, "Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method," *Appl. Opt.* **54**, 1872–1889 (2015).
4. A. Taori et al., "A new method to derive middle atmospheric temperature profiles using a combination of Rayleigh lidar and O_2 airglow temperatures measurements," *Ann. Geophys.* **30**, 27–32 (2012).
5. P. S. Argall, "Upper altitude limit for Rayleigh lidar," *Ann. Geophys.* **25**, 19–25 (2007).
6. P. Keckhut, A. Hauchecorne, and M. L. Chanin, "A critical review of the database acquired for the long-term surveillance of the middle atmosphere by the French Rayleigh lidars," *J. Atmos. Oceanic Technol.* **10**, 850–867 (1993).
7. P. S. Argall et al., "Lidar measurements taken with a large-aperture liquid mirror. 2. Sodium resonance-fluorescence system," *Appl. Opt.* **39**, 2393–2400 (2000).
8. P. S. Argall and R. J. Sica, "A comparison of Rayleigh and sodium lidar temperature climatologies," *Ann. Geophys.* **25**, 27–35 (2007).
9. M. Alpers et al., "Temperature lidar measurements from 1 to 105 km altitude using resonance, Rayleigh, and rotational Raman scattering," *Atmos. Chem. Phys.* **4**, 793–800 (2004).
10. X. Chu and G. C. Papen, "Resonance fluorescence lidar for measurements of the middle and upper atmosphere," Chapter 5 in *Laser Remote Sensing*, T. Fujii and T. Fukuchi, Eds., pp. 191–321, CRC Press, Boca Raton (2005).
11. X. Chu et al., "Fe Boltzmann temperature lidar: design, error analysis, and initial results at the North and South Poles," *Appl. Opt.* **41**, 4400–4410 (2002).
12. X. Chu et al., "First lidar observations of polar mesospheric clouds and Fe temperatures at McMurdo (77.8 degrees S, 166.7 degrees E) Antarctica," *Geophys. Res. Lett.* **38**, L16810 (2011).
13. R. Wing et al., "Atmospheric density and temperature vertical profile retrieval for flight-tests with a Rayleigh lidar on-board the French advanced test range ship Monge," *Atmosphere* **11**, 75 (2020).

14. B. Kaifler et al., "Demonstration of an iron fluorescence lidar rating at 372 nm wavelength using a newly-developed Nd:YAG laser," *Opt. Lett.* **42**, 2858 (2017).

15. R. Wing et al., "Intercomparisons between lidar and satellite instruments in the middle atmosphere," EGU General Assembly, 2020, https://doi.org/10.5194/egusphere-egu2020-11725.

16. R. Wing et al., "Lidar temperature series in the middle atmosphere as a reference data set-Part 2: assessment of temperature observations from MLS/Aura and SABER/TIMED satellites," *Atmos. Meas. Tech.* **11**, 6703–6717 (2018).

17. S. Migliorini, C. Piccolo, and C. D. Rodgers, "Intercomparison of direct and indirect measurements: Michelson Interferometer for passive atmospheric sounding (MIPAS) versus Sonde ozone profiles," *J. Geophys. Res.* **109**, D19316 (2004).

18. M. A. Grawe, K. T. Chu, and J. J. Makela, "Measurement of atmospheric neutral wind and temperature from Fabry–Perot interferometer data using piloted deconvolution," *Appl. Opt.* **58**, 3685–3695 (2019).

19. F. Rocadenbosch et al., "Lidar inversion of atmospheric backscatter and extinction-to-backscatter ratios by use of a Kalman filter," *Appl. Opt.* **38**, 3175–3189 (1999).

20. H.-T. Fang and D.-S. Huang, "Noise reduction in Lidar signal based on discrete wavelet transform," *Opt. Commun.* **233**, 67–76 (2004).

21. J. Mao, "Noise reduction for lidar returns using local threshold wavelet analysis," *Opt. Quantum Electron.* **43**, 59–68 (2012).

22. Z. Zhou et al., "Improvement of the signal to noise ratio of Lidar echo signal based on wavelet de-noising technique," *Opt. Lasers Eng.* **51**, 961–966 (2013).

23. P. Tian et al., "Improved empirical mode decomposition based denoising method for lidar signals," *Opt. Commun.* **325**, 54–59 (2014).

24. M. Sarvani, K. Raghunath, and S. V. B. Rao, "Lidar signal denoising methods-application to NARL Rayleigh lidar," *J. Opt.* **44**, 164–171 (2015).

25. K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.* **62**, 531–544 (2014).

26. H. Li et al., "Efficient lidar signal denoising algorithm using variational mode decomposition combined with a whale optimization algorithm," *Remote Sens.* **11**(2), 126 (2019).

27. B. Sun, H. Fang, and D. Huang, "Lidar signal denoising using least-squares support vector machine," *IEEE Signal Process. Lett.* **12**, 101–104 (2005).

28. H.-T. Fang, D.-S. Huang, and Y.-H. Wu, "Antinoise approximation of the lidar signal with wavelet neural networks," *Appl. Opt.* **44**(6), 1077–1083 (2005).

29. Z. Zhou et al., "Improvement of the signal to noise ratio of Lidar echo signal based on wavelet de-noising technique," *Opt. Lasers Eng.* **51**, 961–966 (2013).

30. A. Haar, "Zur theorie der orthogonalen funktionensysteme," *Math. Ann.* **69**, 331–371 (1910).

31. S. Mallat, *A Wavelet Tour of Signal Processing—The Sparse Way*, Elsevier, Burlington (2009).

32. Z. Gao et al., "Laser range data denoising via adaptive and robust dictionary learning," *IEEE Geosci. Remote Sens. Lett.* **12**, 1750–1754 (2015).

33. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006).

34. R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," Technical Report, Technion, Israel Institute of Technology (2008).

35. H. Wang et al., "Batch algorithms of matching pursuit and orthogonal matching pursuit with applications to compressed sensing," in *Int. Conf. Inf. and Autom.*, Zhuhai, Macau, pp. 824–829 (2009).

36. V. S. Kumar, P. B. Rao, and M. Krishnaiah, "Lidar measurements of stratosphere-mesosphere thermal structure at a low latitude: comparison with satellite data and models," *J. Geophys. Res.* **108**(D11), 4342 (2003).

37. G. Denevi, S. Garbarino, and A. Sorrentino, "Iterative algorithms for a non-linear inverse problem in atmospheric lidar," *Inverse Prob.* **33**(8), 085010 (2017).

38. Joint Committee for Guides in Metrology, "Evaluation of measurement data, supplement 1 to the guide to the expression of uncertainty in measurement propagation of distributions using a Monte Carlo method," Technical report-a (2008).
39. Joint Committee for Guides in Metrology, "Evaluation of measurement data: guide to the expression of uncertainty in measurement," Technical report-b (2008).
40. A. S. Jursa, *Handbook of Geophysics and the Space Environment*, 4th ed., Air Force Geophysics Laboratory, Air Force Systems Command, United States Air Force, Hanscom Air Force Base, Massachusetts (1985).
41. T. Leblanc et al., "Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms—Part 1: vertical resolution," *Atmos. Meas. Tech.* **9**, 4029–4049 (2016).
42. T. Leblanc et al., "Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms—Part 2: ozone DIAL uncertainty budget," *Atmos. Meas. Tech.* **9**, 4051–4078 (2016).
43. T. Leblanc et al., "Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms—Part 3: temperature uncertainty budget," *Atmos. Meas. Tech.* **9**, 4079–4101 (2016).

Varanasi Satya Sreekanth is a research fellow at the National Atmospheric Research Laboratory, Deparament of Space, India. He is a PhD candidate in the Department of Avionics at the Indian Institute of Space Science and Technology, Thiruvananthapuram, India. He received his Master of Technology with distinction in microwave and communication engineering and his Bachelor of Technology in electronics and communication engineering from Jawaharlal Nehru Technological University, Kakinada. He is a recipient of the student research fellowship from the National Atmospheric Research Laboratory, Department of Space, India. His current research interests include lidars, instrumentation, remote sensing, signal processing, and applied machine learning for remote sensning.

Karnam Raghunath is a scientist/engineer at the National Atmospheric Research Laboratory, Department of Space, Government of India, Tirupati, India, in the field of optical remote sensing with atmospheric lidars. He received his bachelor of technology degree in electronics and communication engineering from Nagarjuna University, India, and his master of engineering in communication systems from Regional Engineering College, Trichy, India. He also received his Doctor of Philosophy in electronics and communication engineering from SV University, India. He worked in the field of high-power RF systems and antennas related to atmospheric radars at Society for Applied Microwave Electronics Engineering and Research, Mumbai, India. He has published in Indian and international journals.

Deepak Mishra recieved his BE degree in electrical engineering in 2000, his MTech degree in instrumentation in 2003 from Devi Ahilya University Indore, Mishra, and his PhD from IIT Kanpur in the Electrical Engineering Department in 2007. Later, he was a postdoc researcher at the University of Louisville, Kentucky, USA, in the field of signal processing and system neuroscience. After a brief time from 2009 to 2010 as a senior software engineer at CMC limited Hyderabad, he opted to work as an academic faculty member at the Indian Institute of Space Science and Technology Trivandrum in 2010 and continued to work as an associate professor and head in the Department of Avionics. He is responsible for both research and teaching UG and PG students. Moreover, he was the coordinator for the Mtech program in digital signal processing and developed a virtual reality center of excellence during his stay at IIST. He was also awarded the Young Scientist award from System Society of India for his research work in 2012. His research interest includes signal processing, neural networks and deep learning, computer vision and graphics, and image and video processing.