Coherent Doppler Lidar for Boundary Layer Studies and Wind Energy

by

Aditya Choukulkar

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved February 2013 by the
Graduate Supervisory Committee:

Ronald Calhoun, Chair
Eric Kostelich
Huei-Ping Huang
Alex Mahalov
Patrick Phelan

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

This thesis outlines the development of a vector retrieval technique, based on data assimilation, for a coherent Doppler LIDAR (Light Detection and Ranging). A detailed analysis of the Optimal Interpolation (OI) technique for vector retrieval is presented. Through several modifications to the OI technique, it is shown that the modified technique results in significant improvement in velocity retrieval accuracy. These modifications include changes to innovation covariance portioning, covariance binning, and analysis increment calculation. It is observed that the modified technique is able to make retrievals with better accuracy, preserves local information better, and compares well with tower measurements.

In order to study the error of representativeness and vector retrieval error, a lidar simulator was constructed. Using the lidar simulator a thorough sensitivity analysis of the lidar measurement process and vector retrieval is carried out. The error of representativeness as a function of scales of motion and sensitivity of vector retrieval to look angle is quantified.

Using the modified OI technique, study of nocturnal flow in Owens' Valley, CA was carried out to identify and understand uncharacteristic events on the night of March 27th 2006. Observations from 1030 UTC to 1230 UTC (0230 hr local time to 0430 hr local time) on March 27 2006 are presented. Lidar observations show complex and uncharacteristic flows such as sudden bursts of westerly cross-valley wind mixing with the dominant up-valley wind. Model results from Coupled Ocean/Atmosphere

Mesoscale Prediction System (COAMPS®) and other in-situ instrumentations are used to corroborate and complement these observations.

The modified OI technique is used to identify uncharacteristic and extreme flow events at a wind development site. Estimates of turbulence and shear from this technique are compared to tower measurements. A formulation for equivalent wind speed in the presence of variations in wind speed and direction, combined with shear is developed and used to determine wind energy content in presence of turbulence.

Dedicated to my Father, Prakash Choukulkar

ACKNOWLEDGMENT

I would like to express my deep sense of gratitude for Dr. Ronald Calhoun, my thesis advisor and mentor, for his incredible support, invaluable advice and patience throughout the course of my research. Working with Dr. Calhoun was a rewarding experience that not only benefited me academically, but also helped me develop well rounded life skills that I could not have learned otherwise. I would also like to thank my committee, Dr. Huei-Ping Huang, Dr. Alex Mahalov, Dr. Eric Kostelich and Dr. Patrick Phelan for supervising my dissertation and their valuable feedback. Special thanks to Dr. Edward Patton for his valuable inputs and providing LES simulations which were an essential part of the lidar error analysis presented here.

The ASU lidar group acknowledges the support of NSF grant 0522307 (Program Officer: Stephan P. Nelson), Intel Grant for Feed-Forward Feasibility Study, NASA Wake-Vortex grant, DOE grant DE-EE0005379, and NSF award 1160737 (Program Office: Bradley Smull).

I would like to acknowledge the support of my mom and my brother who stood by me at all times. I also thank my mentors: Pastor Sanghoon Yoo, Alok Ghodke and Ronald Lazaro and the amazing family at The Faithful City for making my journey meaningful and enjoyable!

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

1. ROLE OF LIDAR REMOTE SENSING

Observations of the key processes occurring in the atmosphere are essential for wind farm development, boundary layer studies and improvement of weather and climate models. The dearth of high resolution, high precision profiles and 3D volumes of wind data is a significant impediment to progress in understanding and modeling the lower atmosphere for weather forecasting, air quality, and climate modeling. Ground-based, scanning, coherent Doppler LIDAR (Light Detection and Ranging) can play an important role in providing measurements of fluid dynamical processes occurring over a wide range of scales and covering large spatial extents. There have been numerous studies that have made use of Doppler lidar for boundary layer studies (eg: Banta et al., 2002; Lin et al., 2008), wind energy applications (eg: Smith et al., 2006; Banta et al., 2009; Krishnamurthy et al., 2012) mountain meteorology (Hill et al., 2010; Choukulkar et al., 2012a) and dispersion modeling (eg: Collier et al., 2005; Choukulkar et al., 2011). The 3D coherent Doppler lidar is poised to be a potentially transformative tool as a consequence of the growing interest in the research community and the commercial sector for wind farm assessment, identifying and characterizing extreme events and short term prediction (see Hannon et al., 2008; Banta et al., 2010).

1.1 Principle of Operation of a Lidar

The analysis presented in this dissertation is based on measurements made by a Wind Tracer (Lockheed Martin Coherent Technologies) infrared coherent Doppler lidar. The lidar uses a 2 μm eye safe laser beam to make measurements of the wind field (Wind Tracer Systems User Manual). The lidar transmits a pulse approximately 50 – 80 m in

length and 10 cm in diameter. It has a receiver to collect the backscatter from the aerosols present in the air. This backscatter is then compared with the original transmitted signal to analyze the motion and distance of the aerosol particles present on the air. The radial component of the velocity is estimated from the Doppler shift caused by the average aerosol movement over the measurement volume. It is possible to measure accurate estimates of the radial component of the velocity along the beam direction, which is spatially averaged over the pulse length. The sensing volume of the lidar is a pencil shaped beam, 10-30 cm wide and 50 to 80 m long and transmitted roughly 500 times per second. The length of the pulse is the distance over which the data is spatially averaged and is called a range gate (Wind Tracer Systems User Manual).

Lidar has several distinct characteristics compared to RADAR (RAdio Detection And Ranging) in terms of the accuracy and range of scales. Primarily, due to the low spreading of the laser beam, side-lobe difficulties are avoided and high resolution measurements can be obtained close to the ground, potentially, only meters above or even between surface obstacles for typical ranges of up to 5-10 km. Radial resolution is 50-100 m and azimuthal resolution can be as low as a few meters depending on range and the angular speed of the scanner (configurable).

1.2 Vector Retrieval

The ground-based, scanning, coherent Doppler lidar can play an important role in providing measurements of fluid dynamical processes occurring over a wide range of scales and covering large spatial extents. This information needs to be provided in a useful form to the research and applications communities. Extracting the two-

dimensional velocity field from these radial velocity measurements is a central challenge in Doppler lidar analysis. There are many approaches to retrieve the velocity vectors from radial velocity. Usually this requires additional information, supplied by, for example, theoretical considerations or supplemental observations. The simplest technique for vector retrieval is the Velocity Azimuth Display (VAD – see Browning and Wexler 1968). This method assumes horizontal homogeneity of the wind field and may use the least squares approach to determine horizontally averaged velocity vectors with height. Novel dual Doppler techniques which use two or more remote sensors and a least squares processing algorithm have been developed recently for lidar, (see Cahoun et al., 2006, Hill et al., 2010, and Newsom et al., 2008). In Calhoun et al. (2006), intersecting Range-Height Indicator (RHI) scans are used to create "virtual towers", while Hill et al. (2010), use co-planar RHI scans to identify and track vortices in the lee of mountains. Least squares methods are fast, computationally inexpensive, and give spatially and temporally averaged results. More sophisticated methods such as Four Dimensional Variational Analysis (4DVAR) are also available which integrate the knowledge from the observations into a dynamical model governing the evolution of atmosphere with time (Newsom et al., 2004). The use of dynamical models to carry the information forward in time is computationally intensive and requires solutions to adjoint formulations (Xia et al., 2008; Lin et al., 2008).

Therefore, in an attempt to develop a method that has the temporal and spatial resolution of variational approaches while being computationally inexpensive, the Optimal Interpolation (OI) technique (Xu et al., 2007) was adapted to lidar data (Kongara et al., 2012; Choukulkar et al., 2012a). In this work the background and development of

the optimal interpolation technique (Xu et al., 2007) adapted to work on lidar data and the subsequent modifications to that technique (Choukulkar et al., 2012b) will be presented. The basis for this technique is founded in statistical data assimilation (see Daley 1991) and is widely used in data assimilation techniques for several numerical weather prediction models.

This dissertation will outline the formulation of the optimal interpolation technique presented by Xu et al., (2007), and its application to mountain meteorology and wind energy using a coherent Doppler lidar. A theoretical background of the technique will be presented in Chapter 2. Chapter 3 will describe the modifications made to the technique along with validation of the improvements. Chapter 4 will provide a detailed analysis of lidar measurement errors and possible remedies. Chapter 5 and Chapter 6 will describe application of this technique to studying mountain meteorology and impact of turbulence on wind energy resource characterization.

2. THE OPTIMAL INTERPOLATION TECHNIQUE

2.1 Introduction

Data assimilation is a technique of using the observation data in combination with an atmospheric physics model to arrive at a physically consistent representation of the atmospheric state (Bouttier and Courtier, 2001). The use of data assimilation techniques for weather forecasting can be traced to Panofsky (1949) who used polynomial expansions to fit observations to the analysis domain. This approach was improved upon by Gilchrist and Cressman (1954) who introduced the concept of "region of influence" based on which weights are assigned to observations such that an observation carries maximum weight close to the point of observation and decrease as we move away from it. In addition, Gilchrist and Cressman (1954) also proposed that knowledge of the previous numerical forecast could be used to better estimate the analysis. This information from the previous forecast is now called the "background" or "a priori estimate". Bergthorsson and Doos (1955) then proposed a technique to produce "observation increments" by subtracting the background field from each observation and using these observation increments to produce "analysis increments" which were then added to the background field. This work was the foundation for future developments of the statistical interpolation technique (Gandin 1963; Daley 1991).

Data assimilation has been used for various atmospheric data analysis such as mesoscale models, ocean circulation models, satellite data retrievals etc. Based on the application, data assimilation can be classified into two basic approaches: (a) Sequential data assimilation and (b) Retrospective assimilation or non-sequential assimilation.

Sequential data assimilation makes use of observations made in the past till the time of the analysis (which is the basis of real-time data assimilation) while retrospective assimilation even uses observations made in the future for analysis ex: generation of reanalysis data (Lorenc 1986; Daley 1991). Based on the need, data assimilation techniques can range from relatively simple methods such as Interpolation of Observations or Method of Successive Corrections to more complex methods that march the analysis forward in time such as Four Dimensional Variational Analysis – 4DVAR (Daley 1991).

Statistical interpolation, usually referred to as optimal interpolation (OI) is a data assimilation technique where the a priori information ("background") about the atmospheric state is used in combination with observations to make a maximum likelihood estimate of the atmospheric state called the "analysis" (Daley 1991). In most cases, determining the analysis is an under-determined problem due to scarcity of observations and the analysis not being directly related to the model variables. Lorenc (1986) provides an excellent review of the indeterminacy and other constraints in data assimilation. In the case of lidar data, we need to estimate the 'u' and 'v' components of the wind field from only radial velocity measurements and this leads to the indeterminacy in the analysis. To overcome this under-determinacy, additional information is provided through a priori estimate of the 'u' and 'v' values which acts as the background.

The OI technique takes into account the possibility of errors in both the background and the observations when estimating the analysis. In other words, all pieces of data are given importance based on the level of confidence on their accuracy. This

confidence is determined by estimating the error statistics of the data based on Bayesian probabilistic model (Lorenc 1995). Therefore, the OI technique determines the analysis field that minimizes the variance of the analysis error (Daley 1991). The mathematical formulation of this optimization problem is explained below.

2.2 Mathematical Formulation

In order to determine the confidence in the observation and the background information, we need to compare them to reality. This concept of reality is called "truth" and we denote the state vector of the truth as $\mathbf{x}_t$. The state vector of the background is denoted by $\mathbf{x}_b$, observation by $\mathbf{y}_o$ and finally the analysis is denoted as $\mathbf{x}$. The optimization problem we have to solve is to determine an optimal increment, $\delta\mathbf{x}$ and add it to the background such that the analysis vector is given by:

$$\mathbf{x} = \mathbf{x}_b + \delta\mathbf{x} \tag{1}$$

The analysis increment is estimated based on the discrepancy between the observation and the background. This discrepancy is termed as "innovation" and is denoted by $\mathbf{d}$ such that:

$$\mathbf{d} = \mathbf{y}_o - \mathbf{x}_b \tag{2}$$

The analysis can now be defined in terms of innovation as:

$$\mathbf{x} = \mathbf{x}_b + W * (\mathbf{y}_o - \mathbf{x}_b) \tag{3}$$

where, W is a weighing function that is based on the confidence level in the observation and background.

2.2.1 Modeling the Errors

The presence of errors in the background and the observations is represented in terms of error covariances. Given a background state $\mathbf{x}_b$, the error that separates it from the truth is given by

$$\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t \tag{4}$$

It is assumed that given a large number of realizations, the statistics of the background errors are Gaussian in nature and given by the probability density function (Lorenc 1995):

$$p(\mathbf{x}) = \left((2\pi)^2 |\mathbf{B}|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right) \tag{5}$$

where **B** is the background error covariance matrix given by (Daley 1991):

$$\boldsymbol{B} = \overline{(\boldsymbol{\varepsilon}_b - \bar{\boldsymbol{\varepsilon}}_b)(\boldsymbol{\varepsilon}_b - \bar{\boldsymbol{\varepsilon}}_b)^T} \tag{6}$$

Similarly, the measuring instrument will also have certain error characteristics associated with it. Every instrument has a certain accuracy with which it can make measurements. The error in the instrument reading and the "true" value of the atmosphere is called instrument error. Another type of error which is difficult to quantify is the error of representativeness. Every instrument is able to capture the information about the atmospheric processes only over a certain range of scales; the rest of the information is either lost or misrepresented. For example, a lidar is able to measure processes occurring over scales of 100 m to approximately 1 km. It either cannot measure processes occurring beyond this range of scales or measures it partly resulting in misrepresentation

of the atmospheric state. The error due to this misrepresentation is called error of representativeness. In addition, the observations might not always be located on the analysis grid. This requires that the observation value be interpolated onto the analysis grid. The error induced due to this interpolation is called interpolation error. Therefore, the total error in observation is given by:

$$\boldsymbol{\varepsilon}_{\mathrm{o}} = \mathbf{y}_{\mathrm{o}} - \mathrm{H}(\mathbf{x}_{\mathrm{t}}) \tag{7}$$

which includes instrument error, error in the interpolation operator H and representativeness error. The statistics of the observation errors are also assumed Gaussian and given by the probability density function (Lorenc 1995):

$$p(\mathbf{y}_{\mathrm{o}}|\mathbf{x}) = (2\pi\mathbf{R})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{y}_{\mathrm{o}} - \mathrm{H}(\mathbf{x}))^{\mathrm{T}} \mathbf{R}(\mathbf{y}_{\mathrm{o}} - \mathrm{H}(\mathbf{x}))\right) \tag{8}$$

where **R** is the observation error covariance matrix and is given as (Daley 1991):

$$\mathbf{R} = \overline{(\boldsymbol{\varepsilon}_{\mathrm{o}} - \bar{\boldsymbol{\varepsilon}}_{\mathrm{o}})(\boldsymbol{\varepsilon}_{\mathrm{o}} - \bar{\boldsymbol{\varepsilon}}_{\mathrm{o}})^{\mathrm{T}}} \tag{9}$$

2.2.2 Bayesian Formulation

Using Bayesian analysis we can define the accuracy of our knowledge based on past information (the background) and adjust these probabilities based on new information (the observations). Therefore if P(A) is the probability that A occurred in the past, P(B) is the probability of B occurring and P(B|A) is the conditional probability of B given A has occurred, then according to Bayes theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{10}$$

Rewriting the Bayes theorem in terms of the probability density functions of the errors we get (Lorenc 1995):

$$p(\mathbf{x}|\mathbf{y}_o) = \frac{p(\mathbf{y}_o|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y}_o)} \tag{11}$$

where $p(\mathbf{x}|\mathbf{y}_o)$ gives the measure of our certainty of $\mathbf{x}$ (the analysis) given our knowledge of $\mathbf{y}_o$ (observations). Substituting the Eq. (5) and (8) into the Bayesian analysis equation, Eq. (11) and taking logarithms of the resulting equation, we get:

$$-\ln\left[p(\mathbf{x}|\mathbf{y}_o)\right] = \frac{1}{2}(\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}) + \frac{1}{2}(\mathbf{y}_o - H(\mathbf{x}))\mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \tag{12}$$

Therefore, the $\mathbf{x}$ that minimizes the following cost function is the maximum likelihood estimate of the atmospheric state (Lorenc 1995).

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}) + \frac{1}{2}(\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \tag{13}$$

2.3 Error Covariances

The error covariances are an important component in describing our level of confidence in the background and observations. In a one-dimensional scalar system, the error covariances reduce to the variance or r.m.s. departure from the mean (Eq. 10):

$$\mathbf{B} = \text{var}\,\overline{(\boldsymbol{\varepsilon}_b - \bar{\boldsymbol{\varepsilon}}_b)^2} \tag{14}$$

While in a multi-dimensional system, the error covariance is a square symmetric matrix with dimension of $n \times n$ if the original state vector has dimension of $n \times 1$. Then the diagonal terms represent variances and the off-diagonal terms are the cross-covariances. The error covariance matrix is a positive definite matrix with non-zero variance values (unless the background/observation is a perfect representation of the truth). The statistics of the error covariances depend on the physical processes occurring in the atmosphere. Since the truth is unknown, the exact values of the error covariance matrices cannot be determined and must be estimated by making some assumptions. Therefore the analysis increments estimated through this technique are rarely optimal and hence this technique is referred to as statistical interpolation (Daley 1991).

2.3.1 Background Error Covariances and Correlations

There is usually a discrepancy between the background values at a given grid point with the observation values. The background error covariances help to determine how the information from the observation increment is spread to the analysis grid points. This is especially important in data sparse areas where the analysis increment is completely determined by the background error covariance. Background error covariances also play a role in information smoothing and ensure that the analysis contains scales that are adequately resolved by the model. Therefore, for an accurate representation of the atmospheric state the background errors should be flow dependent (Courtier et al., 1998). Estimation of the background error covariances is simplified if the horizontal and vertical structure of the errors is assumed to be separable (Daley 1991).

In this study we restrict our analysis to a two-dimensional horizontal domain and thus neglect the vertical covariances. In addition, the background error covariances are assumed to be homogeneous and isotropic (Xu and Gong, 2003). Under this condition, the covariance depends only on the relative displacement of the grid points and not the absolute locations (Daley 1991). In addition, the condition of positive definite background error covariance matrix is essential for it to be invertible.

2.3.2 Observation Error Covariances and Correlations

These errors determine the level of confidence in the observations. Therefore, accurate specification of these errors will require knowledge of instrument characteristics, error of representativeness and errors in interpolation of the observations. In addition, there might be a possible for the instruments to have biases. The observation error covariances are calculated assuming that these biases have been accounted for and removed from the observation values. It is important to remove biases from measurements as these can lead to biases in the analysis increment (Daley 1991).

Observation errors are assumed to be uncorrelated, but this might not always be the case especially if the observations come from the same instrument. In practice this can be avoided by thinning dense data or by supplemental assumptions that allow observation errors to be correlated (Xu et al., 2007). In addition, measurement techniques that require background information to pre-process the observations (such as satellite data) can cause artificial correlations between the observation and the background, referred to as the "incest problem" (Bouttier and Courtier, 2001).

2.3.3 Estimation of Error Covariances

Since the "truth" is unknown, the background and observation errors cannot be determined explicitly and hence are usually specified in a statistical sense. The information about the error covariances can be estimated from the departures of the background from the observations, also known as innovation, using the "observation method" given by Hollingsworth and Lonnberg (1986). Other techniques use the information from analysis departures to estimate the error covariances such as the National Meteorological Center's "NMC method" (Parrish and Derber 1992).

There are several properties of error covariances that need to be kept in mind when defining them (Bouttier and Courtier, 2001):

a. They must form a symmetric positive definite matrix.
b. Correlations must be smooth in a physical sense.
c. Correlations must go to zero for large separations so that observations only have a local effect on the analysis.
d. The correlations should show dependence on direction or location that is physically unjustifiable.
e. They should not lead to physically unreasonable background error variance for any parameter.

2.4 Error Covariances for Doppler lidar

Conventional statistical interpolation techniques assume the error covariances are homogeneous, isotropic and Gaussian distributed (Daley 1991). This is not always true in the case of Doppler lidar. Xu and Gong (2003) reported on the non-isotropic and

nearly homogeneous nature of Doppler radar data. Doppler lidar data can be assumed to show similar characteristics. Based on this assumption, the formulations of error covariance functions presented by Xu and Gong (2003) and Xu et al. (2006) are adopted to work with lidar data.

The lidar error characteristics are expected to be significantly different compared to error characteristics from independent measurement platforms. Conventional statistical interpolation techniques require the observation errors to be uncorrelated with the background errors and not auto-correlated. Lidar observations have range-gate resolution of around 60 – 100 m in the radial direction and adjacent beams having an azimuth spacing of approximately 1 degree. Due to this, lidar observations are expected to be correlated in both the radial and azimuth direction. These deviations from expected behavior can accounted for through supplemental assumptions as will be explained in the following sections.

2.4.1 Radial Velocity Error Covariance Function

The covariance of a vector field in a Cartesian coordinate system is defined by:

$$\mathbf{C}_{\mathrm{vv}} = \left\langle \mathbf{v}_1 \mathbf{v}_2^{\mathrm{T}} \right\rangle \tag{15}$$

where, $\mathbf{v}_{\mathrm{i}} = (u_i , v_i)$ is the velocity vector field and at $\mathbf{x}_i = (x_i , y_i)$ and $i = 1, 2$. In our case, the analysis will be performed directly on the lidar scan surface. This not only eliminates the errors due to interpolation, but also simplifies the definition of the covariance functions (Xu et al., 2006). Therefore, a coordinate transformation needs to be done to rotate the Cartesian coordinate to align with the lidar scan surface (see Figure 1 of Xu and

Gong 2003). Taking the transformation of the coordinate system into account, the covariance tensor can now be defined as:

$$\mathbf{C}_{vv} = \mathbf{R}(\alpha)\mathbf{C}(r)\mathbf{R}^{T}(\alpha) \tag{16}$$

where, $\mathbf{R}(\alpha)$ is the rotation matrix that rotates the Cartesian coordinate system to the lidar surface. The resulting velocity vector on the rotated coordinate is denoted by:

$$\left(l_i, t_i\right)^{T} = \mathbf{R}(\alpha)\mathbf{v}_i \tag{17}$$

and its respective covariance tensor by:

$$\mathbf{C}(r) = \left\langle (l_1, t_1)(l_2, t_2)^{T} \right\rangle \tag{18}$$

where, r is the horizontal distance between the two points. The components of the covariance function in Eq. (16) are as follows (Xu and Gong, 2003):

$$\left.\begin{aligned} \langle u_1 u_2 \rangle &= C_{ll}(r)\cos^2\alpha + C_{tt}(r)\sin^2\alpha \\ \langle v_1 v_2 \rangle &= C_{ll}(r)\sin^2\alpha + C_{tt}(r)\cos^2\alpha \\ \langle u_1 v_2 \rangle &= \langle v_1 u_2 \rangle = \{C_{ll}(r) - C_{tt}(r)\}\sin\alpha\cos\alpha \end{aligned}\right\} \tag{19}$$

where, $C_{ll}(r)$, $C_{lt}(r)$, $C_{tl}(r)$ and $C_{tt}(r)$ are the components of $\mathbf{C}(r)$ and $C_{lt}(r) = C_{tl}(r) = 0$ assuming isotropy of the covariance tensor (Xu and Gong 2003; Daley 1991).

Lidar measures only the radial component of the wind vector. Therefore the error in the radial velocity is given by:

$$v_{ri} = u_i \cos\beta_i + v_i \sin\beta_i \tag{20}$$

where, β is the azimuth angle measured clockwise from north. Therefore, the radial velocity error covariance function based on the radial velocities measured by the lidar can be derived using Eqs. (19) and (20) as follows (Xu and Gong, 2003):

$$\begin{aligned}\mathbf{C}_{vr}(\mathbf{x}_1,\mathbf{x}_2) &= \langle v_{r1} v_{r2} \rangle \\ &= \frac{1}{2}\{C_{ll}(r)+C_{tt}(r)\}\cos(\beta_1-\beta_2) \\ &\quad + \frac{1}{2}\{C_{ll}(r)-C_{tt}(r)\}\cos(\beta_1+\beta_2-2\alpha)\end{aligned} \tag{21}$$

In a similar manner, the covariance function of the tangential velocity error can be derived by using the tangential velocity equation given by:

$$v_{\tau i} = -u_i \sin\beta_i + v_i \cos\beta_i \tag{22}$$

Using Eqs. (19), (20) and (22), the tangential velocity covariance function is defined as:

$$\begin{aligned}\mathbf{C}_{r\tau}(\mathbf{x}_1,\mathbf{x}_2) = \mathbf{C}_{\tau r}(\mathbf{x}_2,\mathbf{x}_1) &= \langle v_{r1} v_{\tau 2} \rangle \\ &= \frac{1}{2}\{C_{ll}(r)+C_{tt}(r)\}\sin(\beta_1-\beta_2) \\ &\quad - \frac{1}{2}\{C_{ll}(r)-C_{tt}(r)\}\sin(\beta_1+\beta_2-2\alpha)\end{aligned} \tag{23}$$

The radial velocity error covariance functions given by Eqs. (21) and (23) can be re-written in the following manner (Xu et al., 2007):

$$\mathbf{C}_{rr}(\mathbf{x}_i,\mathbf{x}_j) = \left(\frac{C_+(r)\cos\beta_- + C_-(r)\cos\beta_+}{2}\right) \tag{24}$$

$$\mathbf{C}_{\tau r}(\mathbf{x}_i,\mathbf{x}_j) = \left(\frac{C_+(r)\sin\beta_- - C_-(r)\sin\beta_+}{2}\right) \tag{25}$$

where, $C_+(r) = C_{ll}(r) + C_{tt}(r)$, $C_-(r) = C_{ll}(r) - C_{tt}(r)$, $\beta_+ = \beta_1 + \beta_2 - 2\alpha$ and $\beta_- = \beta_1 - \beta_2$

Eqs. (24) and (25), give the formulation for the radial and tangential velocity error covariances. However, to use these formulations the radial and tangential velocity errors must be known. Since we cannot directly determine the errors in the radial and tangential components, we have use indirect means to estimate them. The next section covers how these error covariance functions are formulated.

2.4.2 Spectral Formulation of the Error Covariance Function

Our observation of the atmosphere is limited to the range of scales that are resolved by our measurement platforms. Therefore our observations can be interpreted as spectrally truncated or smoothed versions of the "truth" (Daley 1991). Therefore, while defining the error covariance functions, it is common to use truncated spectral expansions to define the shape of the error covariance functions (Hollingsworth and Lonnberg, 1986; Xu and Wei, 2001). Xu and Wei (2001) present a formulation for the error covariance functions using truncated spectral expansions. For this analysis, apart from the assumption of homogeneity and isotropy, the structure of the error covariances is assumed to be sufficiently smooth for the covariance functions to be defined using spectral expansions. With these assumptions, the binned covariance data can be fit using truncated spectral expansions as (Xu and Wei, 2001):

$$C_+(r) = \sum \left[S_{ll}(k_i) + S_{tt}(k_i) \right] J_0(k_i r) \tag{26}$$

$$C_-(r) = \sum \left[S_{ll}(k_i) - S_{tt}(k_i) \right] J_2(k_i r) \tag{27}$$

where, the summation is over i = 0 to N, and N is the truncation number for the spectral expansions; J_0 and J_2 are the zeroth order and second order Bessel functions and k_i (i = 1 to N) are the zeros of the Bessel's function $J_1(k\mathrm{D})$, where D is the maximum range of the domain under study. The $k_0 = 0$ represents the constant term in the spectral expansion, which corresponds to the large scale effects. Appropriate selection of the truncation number can ensure that the high wave number data noise is filtered and the scales that are resolved by the instrument are preserved. In the case of the lidar, the truncation number of N = 20 was found to give the best results.

Therefore, the radial and tangential error covariances can be estimated using Eqs. (26) and (27) if the coefficients of the spectral expansions are known. The coefficients of the spectral expansions are determined by fitting the binned innovation covariance data to Eqs. (26) and (27). The detail of how the binned innovation covariance data are determined and the fitting process is explained in the next section.

2.4.3 Innovation Covariance Partitioning

The spectral formulations of the error covariance functions described in the previous section can be used to estimate the error covariances by estimating the spectral coefficients through fitting with innovation covariance data (Hollingsworth and Lonnberg, 1986; Xu and Wei, 2001). Innovation covariance is defined as $<d_i\, d_j>$. However, this formulation of the error covariances assumes that the background wind errors are homogeneous and isotropic in the horizontal domain. Xu and Gong (2003) showed that lidar background errors are non-isotropic. In addition, it is required that the lidar observation errors are not auto-correlated, which is invalid in case of the lidar,

where data from adjacent beams and range-gates is expected to be correlated (Xu and Gong, 2003). Xu et al. (2007) present a reformulation of the innovation method (Hollingsworth and Lonnberg, 1986) which allows the use of lidar data to build the error covariance functions.

This reformulation of the innovation method makes the following assumptions about the radial velocity observation errors: (a) they are not cross-correlated with background errors and (b) they are not auto-correlated beyond a length scale L. While the first assumption is unchanged from the one required for the conventional innovation method, the second assumption relaxes the requirement for observation errors to be uncorrelated spatially. In the light of these assumptions, the innovation correlation can be written as:

$$\left\langle d_i d_j \right\rangle = \frac{1}{\sigma_{\mathrm{d}}^2} \left\langle \left(\mathbf{y}_i - \mathrm{H}_i(\mathbf{x}_{\mathrm{b}})\right)\left(\mathbf{y}_j - \mathrm{H}_j(\mathbf{x}_{\mathrm{b}})\right)^{\mathrm{T}} \right\rangle \tag{28}$$

where, $\langle d_i d_j \rangle$ is the normalized innovation covariance, $\mathbf{y}_i$ is the observation and $\mathrm{H}_i(\mathbf{x}_{\mathrm{b}})$ is the background at the i-th location. Eq. (28) can be re-written as:

$$\begin{aligned}
\left\langle d_i d_j \right\rangle &= \frac{1}{\sigma_{\mathrm{d}}^2} \left\langle \left[\left(\mathbf{y}_i - \mathrm{H}_i(\mathbf{x}_{\mathrm{t}})\right) + \left(\mathrm{H}_i(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_i(\mathbf{x}_{\mathrm{b}})\right)\right]\left[\left(\mathbf{y}_j - \mathrm{H}_j(\mathbf{x}_{\mathrm{t}})\right) + \left(\mathrm{H}_j(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_j(\mathbf{x}_{\mathrm{b}})\right)\right]^{\mathrm{T}} \right\rangle \\
&= \frac{1}{\sigma_{\mathrm{d}}^2} \left\{ \left\langle \left[\left(\mathbf{y}_i - \mathrm{H}_i(\mathbf{x}_{\mathrm{t}})\right)\left(\mathbf{y}_j - \mathrm{H}_j(\mathbf{x}_{\mathrm{t}})\right)^{\mathrm{T}}\right] \right\rangle \right\} \\
&\quad + \frac{1}{\sigma_{\mathrm{d}}^2} \left\langle \left[\left(\mathrm{H}_i(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_i(\mathbf{x}_{\mathrm{b}})\right)\left(\mathrm{H}_j(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_j(\mathbf{x}_{\mathrm{b}})\right)^{\mathrm{T}}\right] \right\rangle \\
&\quad + \frac{1}{\sigma_{\mathrm{d}}^2} \left\langle \left[\left(\mathbf{y}_i - \mathrm{H}_i(\mathbf{x}_{\mathrm{t}})\right)\left(\mathrm{H}_j(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_j(\mathbf{x}_{\mathrm{b}})\right)^{\mathrm{T}}\right] \right\rangle \\
&\quad + \frac{1}{\sigma_{\mathrm{d}}^2} \left\langle \left[\left(\mathrm{H}_i(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_i(\mathbf{x}_{\mathrm{b}})\right)\left(\mathbf{y}_j - \mathrm{H}_j(\mathbf{x}_{\mathrm{t}})\right)^{\mathrm{T}}\right] \right\rangle
\end{aligned} \tag{29}$$

According to our assumptions, the background and observation errors are not correlated. Hence, the last two terms in Eq. (29) will go to zero, reducing Eq. (29) to:

$$\begin{aligned}\left\langle d_i d_j \right\rangle &= \frac{1}{\sigma_{\mathrm{d}}^2}\left\langle \left[\left(\mathbf{y}_i - \mathrm{H}_i(\mathbf{x}_{\mathrm{t}})\right)\left(\mathbf{y}_j - \mathrm{H}_j(\mathbf{x}_{\mathrm{t}})\right)^{\mathrm{T}}\right] + \left[\left(\mathrm{H}_i(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_i(\mathbf{x}_{\mathrm{b}})\right)\left(\mathrm{H}_j(\mathbf{x}_{\mathrm{t}}) - \mathrm{H}_j(\mathbf{x}_{\mathrm{b}})\right)^{\mathrm{T}}\right]\right\rangle \\ &= \frac{1}{\sigma_{\mathrm{d}}^2}\left(\mathbf{C}_{\mathrm{rr}} + \mathbf{C}_{\mathrm{rr}}^{\mathrm{o}}\right)\end{aligned} \tag{30}$$

where, $\mathbf{C}_{\mathrm{rr}}$ is the background error covariance and $\mathbf{C}_{\mathrm{rr}}^{\mathrm{o}}$ is the lidar radial velocity observation error covariance. As mentioned previously, the lidar observation errors are expected to be correlated for a certain length (two range-gates or beams). Beyond this correlation length (L), the lidar observation errors are expected to rapidly diminish to zero. Therefore, with this in mind, the innovation correlation is split as follows:

$$\left\langle d_i d_j \right\rangle = \begin{cases} \dfrac{\mathbf{C}_{\mathrm{rr}}}{\sigma_{\mathrm{d}}^2} & \text{for } r \geq \mathrm{L} \\ \dfrac{\mathbf{C}_{\mathrm{rr}} + \mathbf{C}_{\mathrm{rr}}^{\mathrm{o}}}{\sigma_{\mathrm{d}}^2} & \text{for } 0 \leq r \leq \mathrm{L} \end{cases} \tag{31}$$

where < > denotes the statistical mean, and $d_i=(V_{\mathrm{rd}i} - \langle V_{\mathrm{rd}i}\rangle)/\sigma_{\mathrm{d}i}$ is the normalized innovation for the i-th observation point, $V_{\mathrm{rd}i} = V_{\mathrm{ro}i} - V_{\mathrm{rb}i}$ is the innovation (difference between the background radial velocity $V_{\mathrm{rb}i}$ and the observed radial velocity $V_{\mathrm{ro}i}$ at the i-th observation point), $(\sigma_{\mathrm{d}i})^2=\langle(V_{\mathrm{rd}i} - \langle V_{\mathrm{rd}i}\rangle)^2\rangle$ is the innovation variance at the i-th observation point of the lidar for each pair of $d_i d_j$ and σ_{d} is the averaged standard deviation of the innovation over all the observations points. Therefore, the covariance portioning shown in Eq. (28) allows the use of lidar data as long as the correlation length (L) is properly selected.

As seen from Eq. (28), for range greater than the correlation length, the innovation covariance is made up of only the background error covariance. Therefore, the binned innovation covariance beyond the correlation length (L), and up to the maximum range of the domain (D), can be used to fit the spectral functions defined in Eqs. (26) and (27) to determine the coefficients of the spectral expansions for the background error covariance. If the background error covariance function is smooth (as per our assumptions), then using these coefficients, the background error covariance function can be determined over the unfitted range (r < L) as well.

Once the background error covariance is determined for the unfitted range (r < L), the difference of the innovation covariance and background error covariance for the unfitted range gives the estimate of the observation error covariance for that range. Now, the coefficients for the spectral expansions of the observation error covariance function can be determined in a similar way to the background error covariance functions. In this case, the functions in Eqs. (26) and (27) will be denoted by C_+^o and C_-^o and will rapidly go to zero as $r \rightarrow$ L. Therefore, the functional forms of C_+^o and C_-^o will be fit to the estimate of observation error covariance in the correlated range ($0 \leq r \leq$ L). Using these coefficients, the observation error covariance for the complete range ($0 \leq r \leq D$) can be determined.

2.4.4 Innovation Covariance Binning

The error covariance functions for the lidar are assumed to be functions of only the horizontal distance between the observation pairs (Xu et al., 2007). In addition, the error covariance for a pair of point also depends on their relative location on the lidar

scan. Xu et al. (2007) proposed a two-step binning scheme for innovation pairs as follows:

(a) the innovation data is binned every ± 0.1 for $|\cos\beta_-| < 0.9$ and $|\cos\beta_+| < 0.9$; every 0.1 for $0.9 \leq |\cos\beta_-| \leq 1$ and $0.9 \leq |\cos\beta_+| \leq 1$;

(b) the innovation data within each of the angle bins above are then binned spatially for a given spatial bin size. For the case of the lidar, a spatial bin size of approximately half the range-gate size was found to be optimal.

Xu et al. (2007) recommend considering a bin only if it contains ~10^4 qualified pairs. In the case of the lidar, this requirement was found to be quite restrictive resulting in loss of large amounts of data. Hence, this requirement was reduced to 100 pairs per bin and found to work very well. In addition, Xu et al. (2007) consider only 11 angle bins ($\cos\beta_-$ & $\cos\beta_+$= 1, 0.8, 0.6, 0.4, 0.2 and $\cos\beta_-$ = 1 & $\cos\beta_+$= – 1). In the case of lidar data, this was found to result in covering only 60% of the data and hence, the other 9 bins were considered as well. This binned innovation covariance data is used to perform the fitting described in the previous section.

2.4.5 Estimating background and observation error variance

The variance of the background and the observation errors can be found in the limit $r \to 0$. Xu et al. (2007) show that in the limit $r \to 0$, the functional forms of the covariance functions are given by:

$$\begin{aligned} C_+(r) &= 2\sigma^2 - ar^2 \\ C_-(r) &= br^2 \end{aligned} \tag{32}$$

where, σ is the background error variance; and a and b are constants. Substituting Eq. (32) into Eq. (24) gives:

$$\frac{\mathbf{C}_{rr}}{\sigma_d^2} = \frac{\sigma}{\sigma_d^2} - \left(a - b\cos\beta_+\right)\frac{r^2}{\sigma_d^2} \tag{33}$$

Least squares fitting Eq. (33) using the binned innovation covariance data obtained over the range L < $r \leq$ D, the three unknowns (σ^2, a, and b) can be estimated.

The observation error variance can be estimated by setting r = 0 in Eq. (31) for the range $0 \leq r \leq$ L. It can be seen that setting r = 0 in Eq. (31) results in:

$$\sigma_d^2 = \sigma^2 + \sigma_o^2 \tag{34}$$

where, σ_o^2 is the observation error variance. When estimating the observation error covariance as explained in Section 2.4.3, the estimated $\mathbf{C}_{rr}/\sigma_d^2$ is subtracted from the innovation covariance in the range $0 \leq r \leq$ L to obtain $\mathbf{C}_{rr}{}^{o}/\sigma_d^2$. This data is then rescaled using the estimate of σ_o^2 to obtain data for observation error correlation which is then fit using the spectral expansion in Eq. (26) and (27).

2.5 The Optimal Interpolation Algorithm

Once the error covariance functions are estimated, they can be used to form the background and observation error covariance matrices. Then using the "representer method" proposed by Xu and Gong (2003), the analysis increment can be estimated. This analysis increment when added to the background value gives the analysis, which is the maximum likelihood estimate of the atmospheric state.

The optimal value of the analysis increment can be obtained by minimizing the cost function given in Eq. (13). The equation written in increment form is given below:

$$J = \Delta \boldsymbol{a}^{\mathrm{T}} \boldsymbol{B}^{-1} \Delta \boldsymbol{a} + \left(\mathrm{H}\Delta \boldsymbol{a} - \boldsymbol{d}\right)^{\mathrm{T}} \boldsymbol{R}^{-1} \left(\mathrm{H}\Delta \boldsymbol{a} - \boldsymbol{d}\right) \tag{35}$$

where, $\Delta\mathbf{a}$ is the analysis increment vector. Since, we define our background and perform our analysis on the lidar scan itself, the innovation reduces to $\mathbf{d} = \mathbf{y} - \mathbf{x}_b$, and Eq. (35) reduces to:

$$J = \Delta \mathbf{a}^{\mathrm{T}} \mathbf{B}^{-1} \Delta \mathbf{a} + \left(\Delta \mathbf{a} - \mathbf{d}\right)^{\mathrm{T}} \mathbf{R}^{-1} \left(\Delta \mathbf{a} - \mathbf{d}\right) \tag{36}$$

Therefore, the $\Delta\mathbf{a}$ that minimizes Eq. (36) will result in the optimal analysis increment. Differentiating Eq. (36) with respect to $\Delta\mathbf{a}$ gives:

$$\Delta \mathbf{a} = \mathbf{B}\left(\mathbf{B} + \mathbf{R}\right)^{-1} \mathbf{d} \tag{37}$$

The matrix $\mathbf{B}$ is often too large to be inverted accurately. Xu and Gong (2003) proposed the "representer method" help avoid the inversion of $\mathbf{B}$. The solution of Eq. (37) is broken down into two steps by introducing an "intermediate vector" $\mathbf{z}$, such that:

$$\left(\mathbf{B} + \mathbf{R}\right)\mathbf{z} = \mathbf{d} \tag{38}$$

$$\Delta \mathbf{a} = \mathbf{B}\mathbf{z} \tag{39}$$

where, $\Delta\mathbf{a}$ has a component in the radial and tangential direction given by:

$$\Delta \mathbf{a}_r = \sum \mathbf{C}_{rr} \mathbf{z} \tag{40}$$

$$\Delta \mathbf{a}_\tau = \sum \mathbf{C}_{\tau\tau} \mathbf{z} \tag{41}$$

The intermediate vector can be estimated using Cholesky decomposition (Xu and Gong, 2003) or through variational methods which employ iterative approaches.

3. THE MODIFIED OPTIMAL INTERPOLATION TECHNIQUE

3.1 Theoretical Background

Data assimilation is a process of importing information from the observations in order to find the model representation most consistent with the observations. Data assimilation techniques attempt to estimate the state of the atmosphere by combining observations with an a priori estimate obtained from either from previous observations or numerical models. The new information is imported in the form of observations, and the output, known as the analysis, is a maximum likelihood estimate of the atmospheric state.

The specification of proper innovation statistics is an important factor in the performance of OI technique. Conventional OI techniques usually make use of innovation (defined as observation minus background) statistics derived from independent instruments, for example, radiosonde, sodars etc depending upon availability and requirement, which are assumed to satisfy the requirements that the background errors are homogeneous and isotropic. In addition, observation errors from independent instrument stations are expected to be uncorrelated. Xu et al. (2007) presented a modified innovation method which allows observation errors to be correlated, thus allowing the use of radar data for assimilation. Nevertheless, independent measurements, either from another instrument or numerical weather prediction model, are still required to form the background and innovation statistics. This approach is not always desirable in cases, for example, where a single lidar performs measurements without the benefit of concurrent numerical forecasting. It was found that the retrievals from the VAD technique serve well as the background (Kongara et al., 2012). The VAD assumes

horizontal homogeneity over a range circle and hence satisfies the requirement that the background wind errors are horizontally homogeneous and isotropic over local scales (Xu et al., 2007; Daley 1991). This enables the use of OI technique for vector retrieval in absence of other independent instruments or data from numerical models.

The formulation of the optimal interpolation technique presented by Xu *et al*. (2007) has been adapted for lidar data in Kongara et al. (2012). This implementation of Xu et al. (2007) is henceforth referred to as KON12. This paper will discuss some of the modifications made to the KON12 technique. While these modifications are incremental improvements, the changes result in significant improvements in the retrieval accuracy.

3.1.1 Innovation Covariance and Partitioning

An innovation is defined as observation minus the background. As mentioned above, background errors are assumed to be homogeneous and isotropic. The following assumptions are made about the radial velocity observation errors (Xu et al., 2007): 1) they are not cross-correlated with background errors, and 2) they are not auto-correlated beyond a certain correlation length (L). With these assumptions in mind, the innovation correlation can be split as follows (Xu et al., 2007):

$$< d_{\mathrm{i}} d_{\mathrm{j}} >= \begin{cases} \dfrac{\mathbf{C}_{\mathrm{rr}}}{\sigma_{\mathrm{d}}^2} & r > \mathrm{L} \\ \dfrac{(\mathbf{C}_{\mathrm{rr}} + \mathbf{C}_{\mathrm{rr}}^{\mathrm{o}})}{\sigma_{\mathrm{d}}^2} & r \le \mathrm{L} \end{cases} \tag{42}$$

where < > denotes the statistical mean, and $d_i=(V_{\mathrm{rd}i} - <V_{\mathrm{rd}i}>)/\sigma_{\mathrm{d}i}$ is the normalized innovation for the i-th observation point, $V_{\mathrm{rd}i} = V_{\mathrm{ro}i} - V_{\mathrm{rb}i}$ is called the

innovation (difference of the background radial velocity V_{rbi} and the observed radial velocity V_{roi} at the i-th observation point), $(\sigma_{di})^2 = <(V_{rdi} - <V_{rdi}>)^2>$ is the innovation variance computed at the i-th observation point of the lidar for each pair of $d_i d_j$ and $(\sigma_d)^2$ is the averaged innovation variance over all the observations points. $\mathbf{C}_{rr}$ is the background error covariance and $\mathbf{C}_{rr}^{o}$ is the lidar radial velocity observation error covariance.

For this relation to be valid, the correlation length needs to be properly selected. It has been observed that the value of correlation length depends on the flow and may vary considerably. For example, for a set of data used in the validation (from Joint Urban 2003 study), over a period of 5 minutes, the correlation length varied between 104 m to 595 m. Therefore, in the modified OI technique, this correlation length is determined in real-time, on a case by case basis, rather than being prescribed to a constant value (Xu et al., 2007). The correlation length is estimated by the plotting the innovation correlations and estimating the range (L) at which the correlation value comes within a specified tolerance value to zero. This method of estimation of the correlation length on a case by case basis ensures the validity of the above partitioning of the innovation correlation.

Determining the correlation length on a case-by-case basis and partitioning the innovation covariance matrix based on that is seen to show improvements in estimating the observation and background error variance as well. Observation error variance plays an important role in the retrieval accuracy. An accurate estimate of the observation error variance is required to re-scale the innovation covariance as explained in Section 2.4.5. And since the background error variance is required to estimate the observation error

variance (see Eq. 34), accurate determination of this parameter is also important. The ability of the modified OI technique to estimate the observation error covariance was tested using an ideal test case.

A uniform and steady flow field of wind flowing from the south-west to the north-east was created. This flow field is considered as our representation of the "truth". Lidar observations of this ideal flow field were created with different values of observation error variance. Vector retrieval was performed on these simulated observations using both KON12 and the modified OI technique. The estimates of the observation error variance from both these techniques are tabulated in Table 1.

Table 1: Skill of OI technique in evaluating the error variance

Parameter	**KON12**		**Modified OI**	
	Supplied	**Estimated**	**Supplied**	**Estimated**
Background error variance	0	0.812	0	0.009
	1	4.551	1	1.067
Observation error variance	0	4.112	0	0
	2.25	23.760	2.25	2.28

As seen from Table 1, the modified OI technique is able to reproduce the actual error variance much more accurately compared to KON12. Accurate estimate of the observation error variance will result in better scaling of the innovation covariance and hence accurate fitting of the spectral functions (see Section 2.4.2).

3.1.2 Innovation Covariance Binning

In Xu *et al*. (2007) and KON12, the innovation covariance is calculated based on assumption of ergodicity. That is, the statistics of the innovation covariance are constant over the given time period. Thus, under the assumption of ergodicity, the statistical mean of $<d_i d_j>$ was computed as the time and spatial mean of $d_i d_j$ (Xu et al., 2007). Here, we propose to relax the assumption of ergodicity. As is observed in our test data set, Doppler lidar data is frequently developing and highly transient over time scales of interest. In order to preserve transient characteristics of the flow and the temporal resolution of the lidar data, the statistics of $d_i d_j$ are calculated from a spatial ensemble at each time step. As a consequence, the background error covariance and the observation error covariance, which are estimated from the innovation covariance (Xu et al., 2007), are also estimated for every time step and preserve the temporal character of the error covariances.

3.1.3 Estimating the Analysis Increment

The analysis increment is estimated using the representer method outlined in Xu et al. (2006). This method consists of a two step procedure to estimate the analysis increment, which is defined by the following expression obtained from minimizing the cost function:

$$\Delta\mathbf{a} = \mathbf{B}\left[\mathbf{B} + \mathbf{R}\mathbf{I}\right]^{-1}\mathbf{d} \tag{43}$$

An intermediate vector **z** is defined as follows in equation (44).

$$\Delta\mathbf{a} = \mathbf{B}\mathbf{z} \tag{44}$$

Substituting equation (44) in equation (43) and simplifying we get (see Xu et al., 2006 for details)

$$\mathbf{z} = \left(\mathbf{C}_{\mathbf{rr}} + \mathbf{C}_{\mathrm{rr}}^{\mathrm{o}}\right)^{-1}\mathbf{d} \tag{45}$$

where *d* is the innovation as defined previously. Here, $\Delta\mathbf{a}$ is the analysis increment and **B** is the background error covariance matrix defined similarly as in Xu et al. (2006).

Instead of solving for the analysis increment through Cholesky decomposition or variational techniques, the following steps are taken to estimate the analysis increment to the *u* and *v* components of the velocity directly. From equation (44) and (45), the analysis increments in the radial and tangential direction can be written as

$$\Delta\mathbf{a}_{\mathrm{rr}} = \mathbf{C}_{\mathrm{rr}}\mathbf{z} = \Delta v_r = \Delta u\cos\beta + \Delta v\sin\beta \tag{46}$$

$$\Delta\mathbf{a}_{\mathrm{tr}} = \mathbf{C}_{\mathrm{tr}}\mathbf{z} = \Delta v_\tau = -\Delta u\sin\beta + \Delta v\cos\beta \tag{47}$$

Here, $\Delta\mathbf{a}_{\mathrm{rr}}$ and $\Delta\mathbf{a}_{\mathrm{tr}}$ are the analysis increments in the radial and tangential directions, Δu and Δv are the analysis increments for and *u* and *v* components of the velocity respectively, β is the azimuth direction measured from the geographic north and $\mathbf{C}_{\mathrm{tr}}$ is the background error cross-covariance in the radial and tangential direction as defined in Xu et al. (2006). Rearranging the terms, we get the following set of equations for Δu and Δv.

$$\Delta u = \Delta\mathbf{a}_{\mathrm{rr}}\cos\beta - \Delta\mathbf{a}_{\mathrm{tr}}\sin\beta \tag{48}$$

$$\Delta v = \Delta\mathbf{a}_{\mathrm{rr}}\sin\beta + \Delta\mathbf{a}_{\mathrm{tr}}\cos\beta \tag{49}$$

Therefore, using equations (48) and (49), the analysis increments in the *u* and *v* velocity components can be estimated directly. This method also has the advantage of not being computationally intensive as there are no complex matrix operations involved.

3.2 Validation

In order to test the effect of the above modifications on retrieval accuracy, the following three step validation is performed. First, the method is tested on idealized flow cases and winds generated by the atmospheric module of the Coupled Ocean/Atmosphere Mesoscale Prediction System – COAMPS® (Hodour, 1997) which are used as the "truth", allowing comparison between the modified OI method and KON12. Next, the algorithms are applied on lidar data from Joint Urban 2003 and the retrievals are tested for their ability to accurately retain local information. Finally, the retrievals from the two techniques are compared to tower measurements.

3.2.1 Ideal test cases

The accuracy of the retrieval algorithm was estimated using idealized test cases. The exact vector field "truth" is therefore known a priori. From the exact fields, radial velocities (as would be observed by a Doppler lidar at the origin of the coordinate system) were calculated by taking the dot product of the velocity vectors with a unit vector along the beam pointing direction. Optimal interpolation is used to retrieve the velocity vectors and which are compared with the exact vectors to estimate the accuracy of the retrieval algorithm. Three test cases were designed. The first test case was a uniform wind, invariant with time, flowing from southwest to the northeast at 5 m s^{-1}. The second case was uniform wind flowing southwest to northeast, but the wind speed

varying with time. The third case was of a convergent flow, time invariant, with winds coming from southwest and southeast. For the first test case, the RMSE from both KON12 and the modified OI technique were zero. This result is expected since the VAD retrieval is used as the background in both cases (and the background is equal to the observation, yielding a zero analysis increment). This test case shows that in the condition that the flow is uniform and non-transient, the two methods converge and give same results. For the second test case, the RMSE of retrieval using KON12 was 0.57 m s^{-1} while the RMSE for the modified OI technique remained zero. From this test, it is seen that estimating the background errors as a function of time improve the accuracy of the retrieval process. The RMSE of the velocity retrieval in the converging flow case was 1.087 m s^{-1} for KON12 and 0.873 m s^{-1} for the modified OI technique. In this test case of spatially inhomogeneous but non-transient flow, the modified technique is able to capture the spatial variations slightly better. This improvement likely results from improved estimation of the correlation length. The results from this study are summarized in Table 2.

Table 2: Validation of retrieval techniques with ideal test cases.

Test Case	**Retrieval Error (KON12)**	**Retrieval Error (Modified OI)**
Test Case I	0 m s^{-1}	0 m s^{-1}
Test Case II	0.57 m s^{-1}	0 m s^{-1}
Test Case III	1.087 m s^{-1}	0.873 m s^{-1}

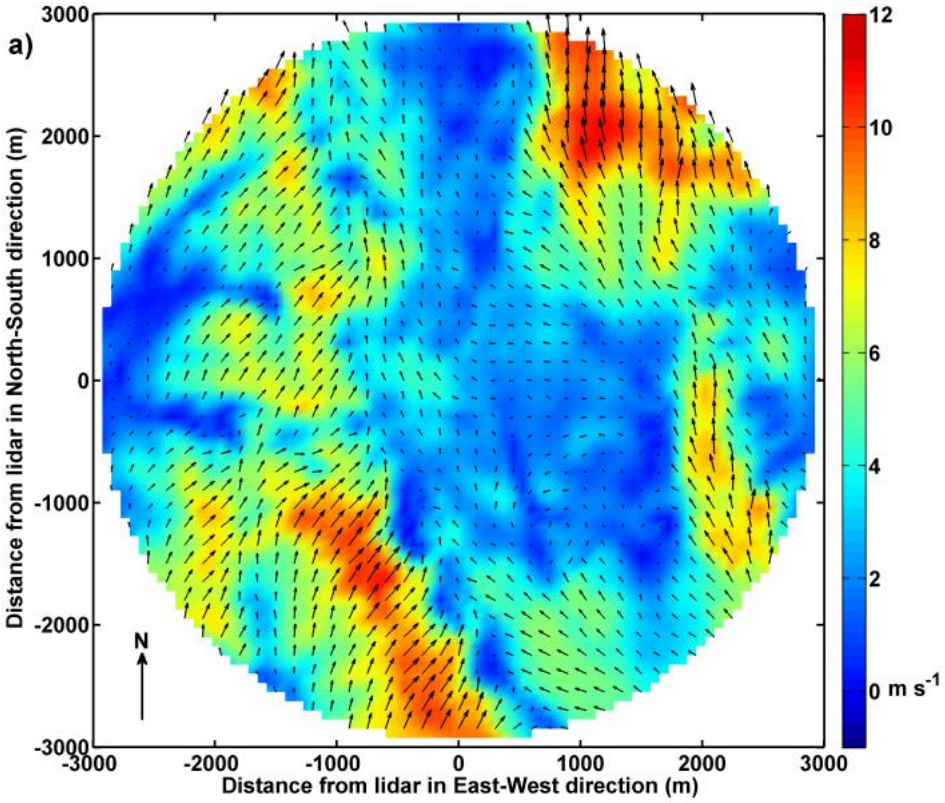
It was also observed that the retrievals are more accurate in the mean wind direction and less so perpendicular to the mean wind direction (Choukulkar et al., 2012a). While the modifications improve accuracy along the mean wind and perpendicular directions, there remains a higher inaccuracy in the perpendicular direction, as expected. For example, the accuracy of the retrieval for the convergent flow case, using the modified OI technique, was approximately 0.45 m s^{-1} in the mean wind direction and 1.14 m s^{-1} in the direction perpendicular to the wind direction. Using KON12, the error came out to be 0.57 m s^{-1} in the mean wind direction and 1.38 m s^{-1} perpendicular to the mean wind direction.

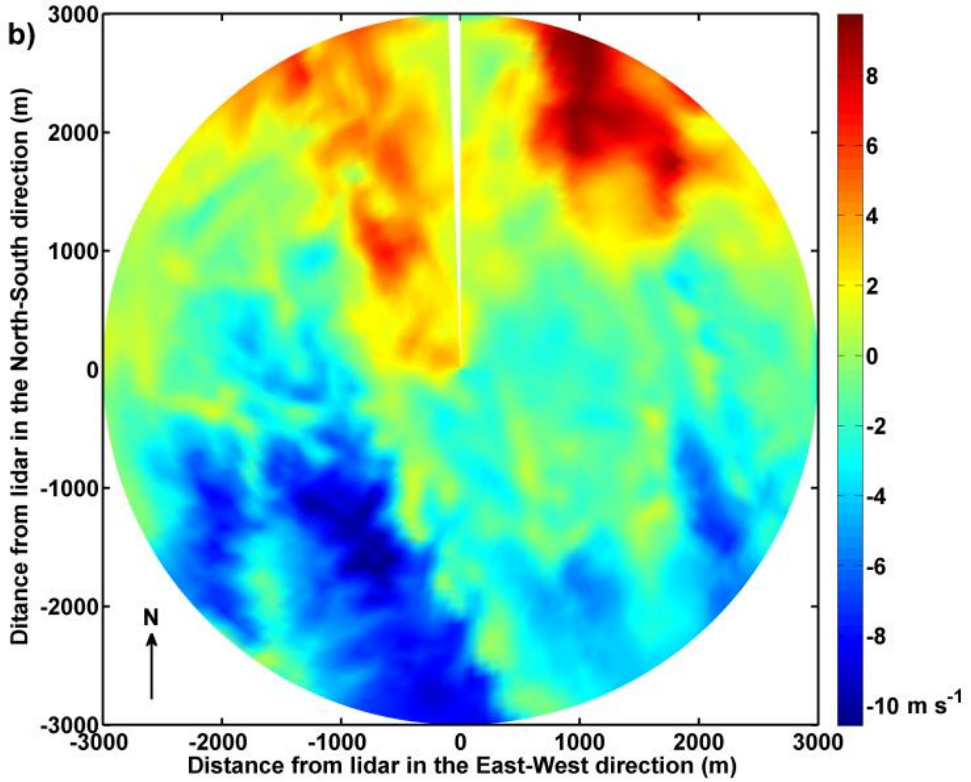
3.2.2 COAMPS simulated flow validation

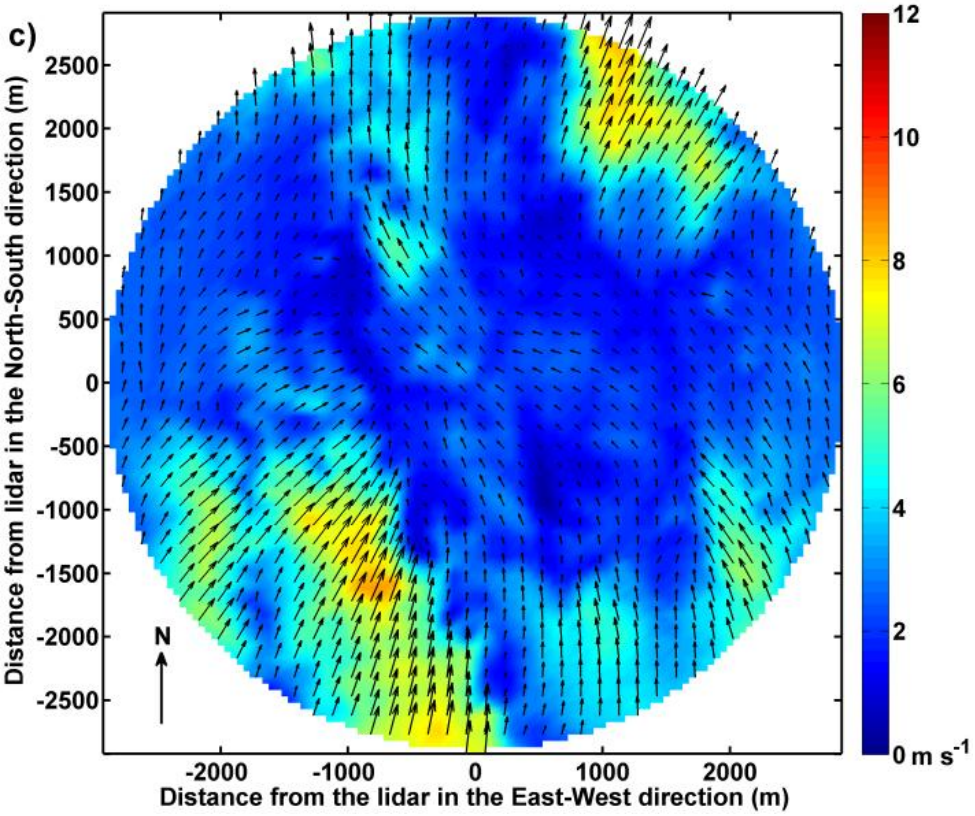
Further tests of the retrieval algorithm accuracy were carried out by treating the wind field from a COAMPS simulation as "truth". Simulation results from a high resolution COAMPS simulation (five grid meshes) are used as the "truth" wind field. The horizontal resolutions of the five grid meshes are 27 km, 9 km, 3 km, 1 km, and 333 m. The domain is centered on Owens Valley, California (USA). The simulation results from the 330-m terrain following height surface on March 27, 1030 UTC were interpolated on to a polar grid. These interpolated vectors are depicted in Figure 1a. Again, the dot product of the "truth" wind vectors with the unit vector in the beam direction provided the radial velocity, as would be measured by a lidar at the center of the coordinate system.

The radial velocity thus obtained from the "truth" is depicted in Figure 1b. Both KON12 and the modified OI algorithm were run with these radial velocities as input.

The retrieved vectors were then compared with the "truth" wind vectors to assess retrieval accuracy. The RMSE of the retrieval was found to be 0.85 m s^{-1} for the modified OI technique compared to the RMSE of 1.81 m s^{-1} for KON12. The retrievals from KON12 and the modified OI technique are presented in Figure 1c and Figure 1d. The RMS of the magnitude of the "truth" wind field was 2.23 m s^{-1}. The flow in this test case, in addition to being spatially inhomogeneous also exhibits considerable temporal variation. This case provides evidence that in more complex cases, calculating time varying error covariances enables to make retrievals with higher degree of accuracy.




b)
Ditance from lidar in the North-South direction (m)
Distance from lidar in the East-West direction (m)
N
3000
2000
1000
0
-1000
-2000
-3000
8
6
4
2
0
-2
-4
-6
-8
-10 m s-1


c)
Distance from lidar in the North-South direction (m)
Distance from the lidar in the East-West direction (m)
N
2500
2000
1500
1000
500
0
-500
-1000
-1500
-2000
-2500
12
10
8
6
4
2
0 m s-1

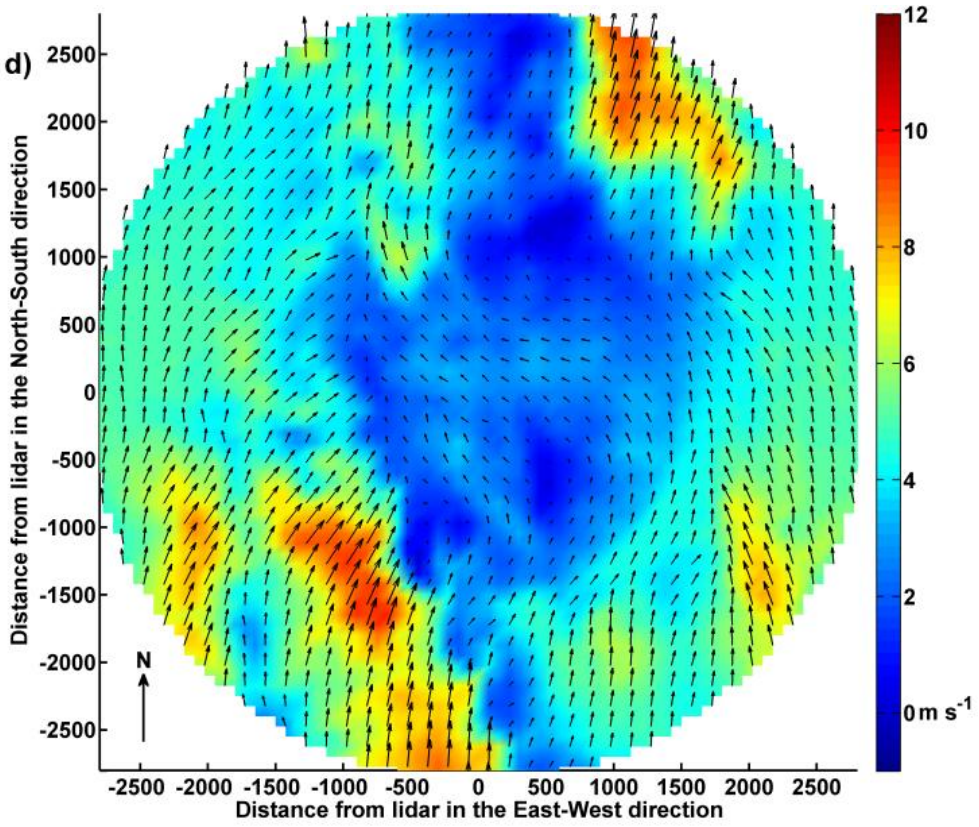


Figure 1. Comparison of retrieved vectors to "truth" vectors. a) "Truth" wind vectors from COAMPS results. b) Radial velocity calculated from "truth" vectors to simulate lidar measurement. c) Vectors retrieved from KON12. d) Vectors retrieved from modified OI technique. Colors represent the magnitude of velocity.

3.2.3 Lidar data analysis

In order to test the ability of the methods to preserve local information, the two techniques were applied to Doppler lidar data gathered on June 27th 2003 during the Joint Urban 2003 experiment (Allwine and Flaherty, 2006). The retrievals from the modified OI technique are compared to those obtained from KON12. In this case, the dot products of the retrieved velocity vectors, (both from KON12 and the modified OI algorithm) were taken along the beam direction to re-obtain the radial velocity. This

"back-calculated" radial velocity was compared with the original lidar retrieved radial velocity to ascertain how well the retrieved information matches the actual measured information. Figure 2 shows the radial velocity comparison between the retrievals from the two techniques for the 12th range-gate on all beams.

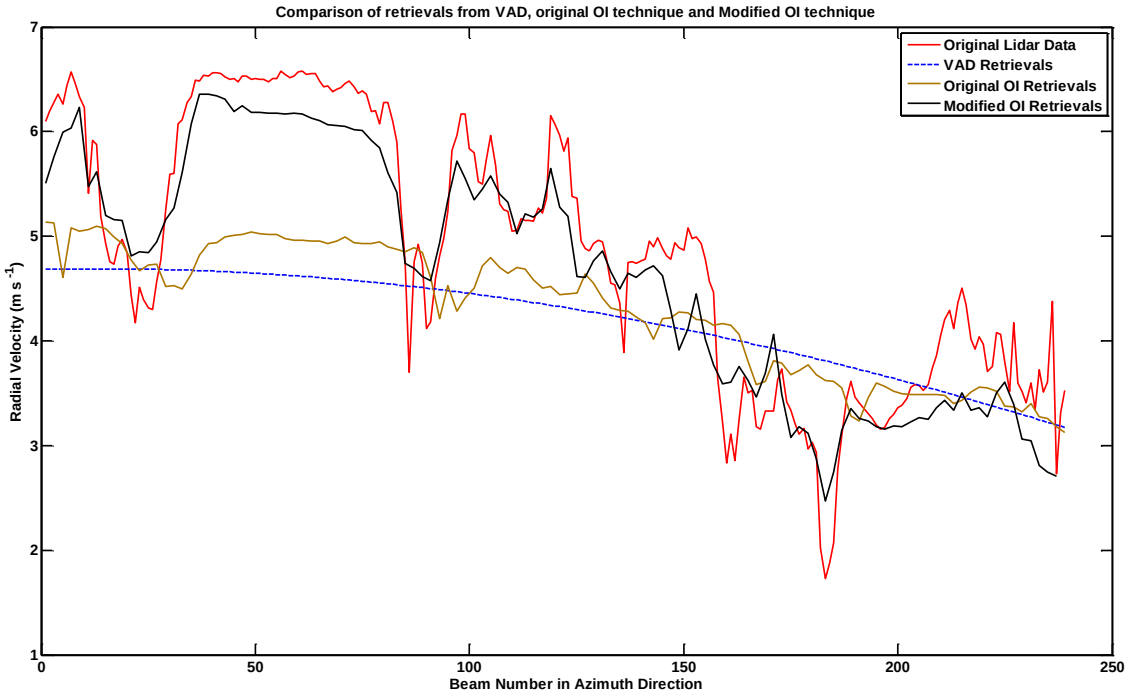


Figure 2. Radial velocity from lidar measurements compared with "back-calculated" radial velocity from retrieval algorithms. Here the radial velocity from the 12th range-gate from all beams is plotted. The x – axis is the beam number in the azimuth direction.

As can be seen from Figure 2, KON12 performs slightly better than the VAD technique and appears to have a time lag with respect to the lidar observations. In addition, it is seen that the retrievals are smoothed. The observed time lag is likely due to the fact that the error statistics are calculated as a temporal average over 5 mins. This temporal averaging also has the additional effect of smoothing the retrievals. The

modified OI technique on the other hand, which calculates the error statistics for every time step, is seen to preserve the local information better than KON12 and VAD almost everywhere.

3.2.4 Comparison with tower data

To further test the validity of the retrievals from the two techniques, the lidar retrievals were compared with tower measurements recorded during a wind farm characterization study in Colorado (USA). Figure 3 shows the comparison of the 10 minute velocity retrieved by the lidar at the tower location using KON12 (Figure 3a) and the modified OI technique (Figure 3b) against the anemometer measurements. It is seen that the RMS difference between the lidar and tower measurements was 2.722 m s^{-1} with KON12 and 1.478 m s^{-1} with the modified OI technique.

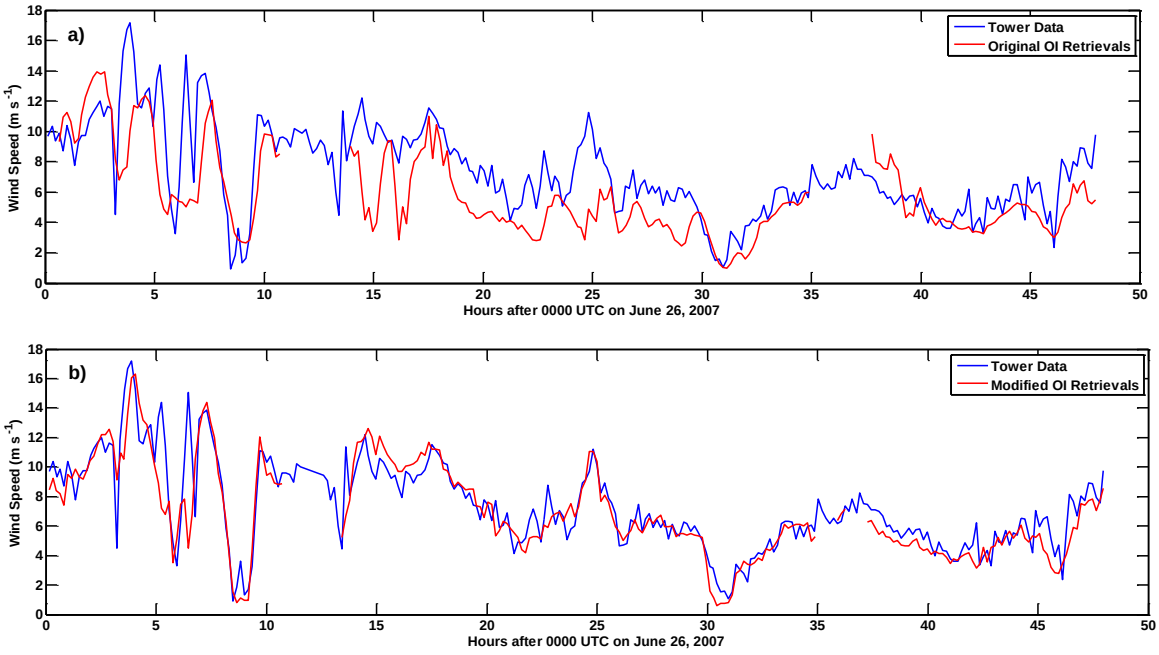


Figure 3. Comparison of OI retrievals with tower measurements. a) KON12 results; b) Modified OI results. The discontinuity in lidar retrievals is due to missing data.

It is to be noted that this technique is sensitive to the number and quality of data points. In KON12 where multiple scans are used to estimate the error covariances, more data points are available compared to the modified technique where only one scan is used. Although this tends of smooth out the results, sometimes it might result in a better analysis (see Figure 3, between 28 and 33 hours). This usually happens when there aren't enough data points at a location from one scan. These occurrences are rare and don't have a significant impact on the accuracy.

3.3 Conclusions

In this paper, simple modifications to KON12, which is an implementation of Xu *et al*. [18], are suggested. The first modification allows calculating the correlation length on a case by case basis and results in determining the innovation covariance partitioning more accurately. The second modification changes the way the innovation covariance is calculated making the error covariances time-varying. The effect of the individual modifications on retrieval accuracy is case dependent. As seen in section 3, the two methods converge when the flow is both uniform and steady. For non-uniform and steady flows, the modified technique of determining the correlation length case by case, results in some improvement in the retrieval accuracy. For more complex flows, it has been observed that by calculating the correlation length on a case by case basis and making the error covariances time-varying, the temporal statistics of the error fields of the flow are retained leading to better preservation of the local flow effects. The combined result of these modifications improves retrieval accuracy significantly while preserving the local information of the flow. In addition, a simple modification to the calculation of the analysis increment, results in reducing the computation time by 20

seconds per scan, bringing the total time to 210 seconds per scan. This modest improvement becomes important considering the huge data analysis needs.

4. CHARACTERIZING LIDAR MEASUREMENT ERRORS

4.1 Introduction

In all measurement processes, errors and uncertainties exist that need to be understood and quantified in order to establish the validity of the results. Here error is defined as the difference between the observed value and the "truth". We divide the error into three components: (1) instrument error, (2) error of representativeness and (3) vector retrieval error. The instrument error is introduced due to non-ideal nature of the instrument such as limited spectral resolution of the detector, frequency chirp in the transmitted laser signal and detector noise among others (see Rye and Hardesty 1993; Dabas et al., 1998; Frehlich 2004). The error of representativeness results from the way the instrument samples the atmosphere. Since the lidar cannot resolve all the scales of motion present in the atmosphere, some scales are not captured. As the true value is unknown there is no direct means of estimating the errors and uncertainty in the measurements.

In the case of the coherent Doppler lidar, there has been considerable work done in understanding the nature of instrument error (Lenschow et al., 2000; Frehlich 2004). However, scant literature is available on sensitivity of lidar measurements to scales of turbulence and look angles. In this chapter, a two part analysis will be carried out that (1) analyzes the sensitivity of lidar measurement to scales of turbulence and look angles and (2) analyzes sensitivity of vector retrieval algorithms to scales of turbulence and look angles. In order to perform this analysis, a lidar simulator was developed that can

simulate lidar measurements in any sufficiently resolved wind field. The details of the lidar simulator are explained in the next section.

4.2 Lidar Simulator

With a lidar simulator it is possible to explore the lidar measurement process and sensitivity analysis of retrieval techniques without the need for expensive field deployments. In addition, it is possible to test lidar measurements for several controlled flow conditions which might not be possible in real deployments. The lidar simulator constructed here uses a flow field generated using Large Eddy Simulation (LES) as the "truth" and simulates lidar measurement of this flow field. With the help of this simulator it is possible to explore the measurement characteristics of a Doppler lidar in different flow regimes. In addition, using the simulated measurements, the sensitivity analysis of a newly developed vector retrieval algorithm for coherent Doppler lidar (see Choukulkar et al., 2012) is presented.

The lidar simulator developed by the lidar group at ASU in collaboration with Dr. Edward Patton at National Center for Atmospheric Research (NCAR) is capable of performing "measurements" at user-defined values of scanning parameters such as Pulse Repetition Frequency (PRF), scanning rate (in degrees per second), pulse averaging and range-gate size. The measurement process is simulated to be close to the operation of a real lidar based on Banakh and Smalikho (1994) and as a result, the radial velocity is a double filtered through the Gaussian pulse intensity and rectangular distance travelled by the pulse during the observation period. This is achieved through the use of the Range-gate Weighing Function (RWF) explained in Dunne et al. (2011).

Another important advantage of the lidar simulator is the ability to perform measurement from multiple locations. This allows analysis of the lidar's measurement accuracy as a function of the look direction. As explained previously the Doppler lidar measures only the component of the wind parallel to the beam angle. As a result, measurements made looking directly into the wind capture a bigger component of the wind velocity and give more accurate retrievals compared to looking orthogonal to the wind speed direction. Through the lidar simulator, it is possible to quantify the degradation in retrieval accuracy as a function of the look angle.

4.3 Setup

In this chapter we will examine in detail two components of the lidar error: (1) error of representativeness and (2) vector retrieval error. The error of representativeness is a function of scanning parameters and scales of motion in the flow. The retrieval error is primarily a function of angle of beam to wind direction (look angle). The LES data simulated wind blowing from the west with an average wind speed of 8 m s^{-1} over a plant canopy. The LES produced results on a 5 km x 5 km grid with a resolution of 2.5 m. The results were produced on horizontal planes at heights 5 m, 10 m, 20 m and 40 m above the ground. For the sake of this analysis, the results on the 40 m level are used. The details of the lidar simulator scanning strategy are explained in the following sections.

4.3.1 Effect of look angle

To characterize the effect of look angle on lidar measurement error, the lidar simulator scans the LES field in three look direction: (1) looking directly into the flow

(east-west orientation), (2) looking perpendicular to the flow (north-south orientation) and (3) looking at the flow with a 45^{o} angle. The original LES field is shown in Figure 4.

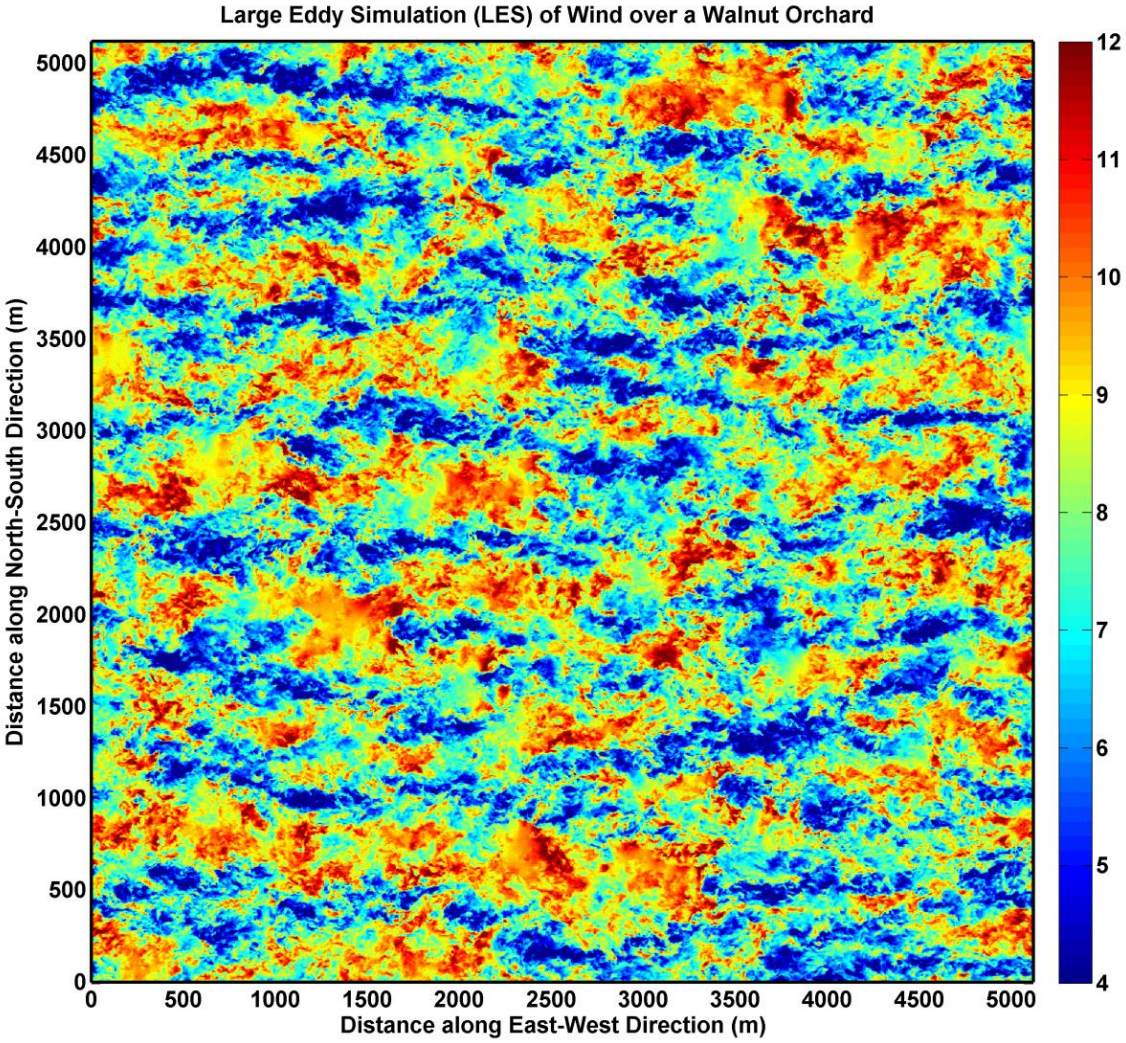
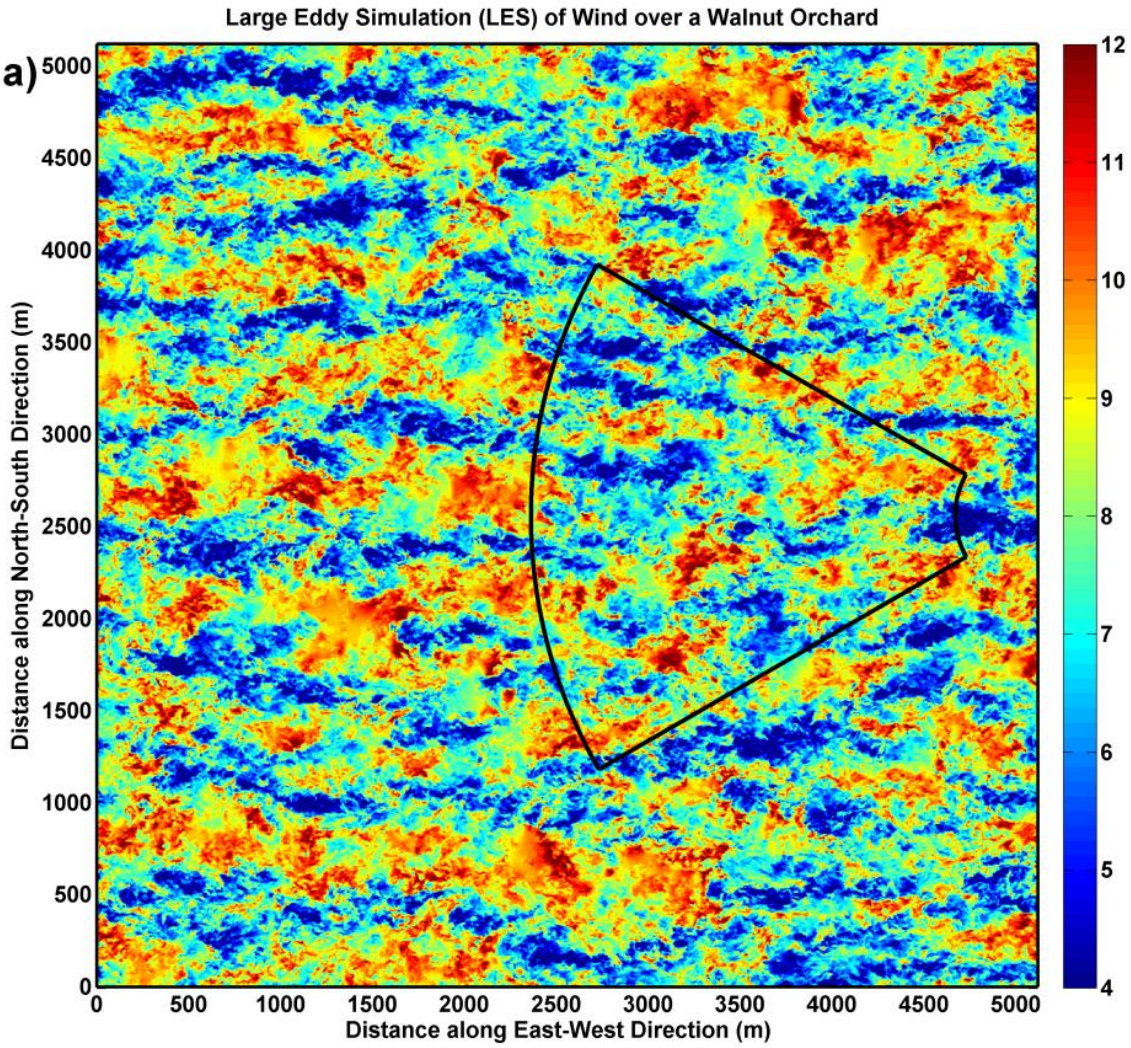
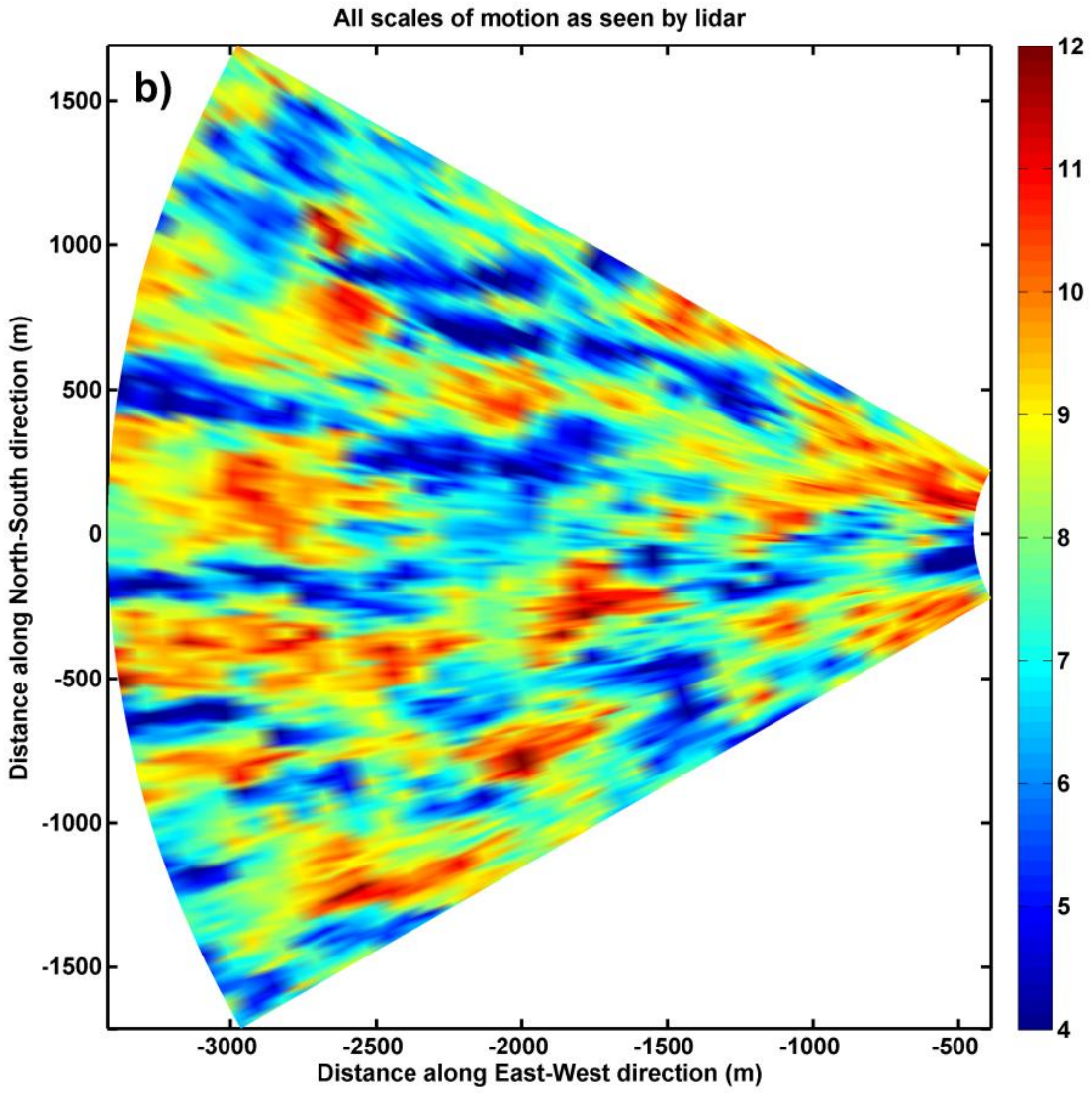


Figure 4. LES result of a flow over a walnut canopy at height of 40 m from the ground. Colors indicate magnitude of wind speed in m s^{-1}.

In this velocity field, lidar measurements were simulated assuming the lidar was positioned at X = 5120 m and Y = 2560 m. The lidar was pointed directly into the wind and scanning from azimuth angle 240^{o} (clockwise from north) to azimuth angle of 300^{o}.

The scanning rate was 2 degrees/second with range-gate size set to 80 m. The simulator was set to produce a Pulse Repetition Frequency (PRF) of 500 per second and averaged 100 pulses to produce 5 Hz data, a typical configuration in our field experiments with a modern long-range Doppler lidar (Wind Tracer). The velocity field as seen by the lidar is reproduced in Figure 5.




All scales of motion as seen by lidar
b)
1500
1000
500
0
-500
-1000
-1500
Distance along North-South direction (m)
-3000
-2500
-2000
-1500
-1000
-500
Distance along East-West direction (m)
12
11
10
9
8
7
6
5
4

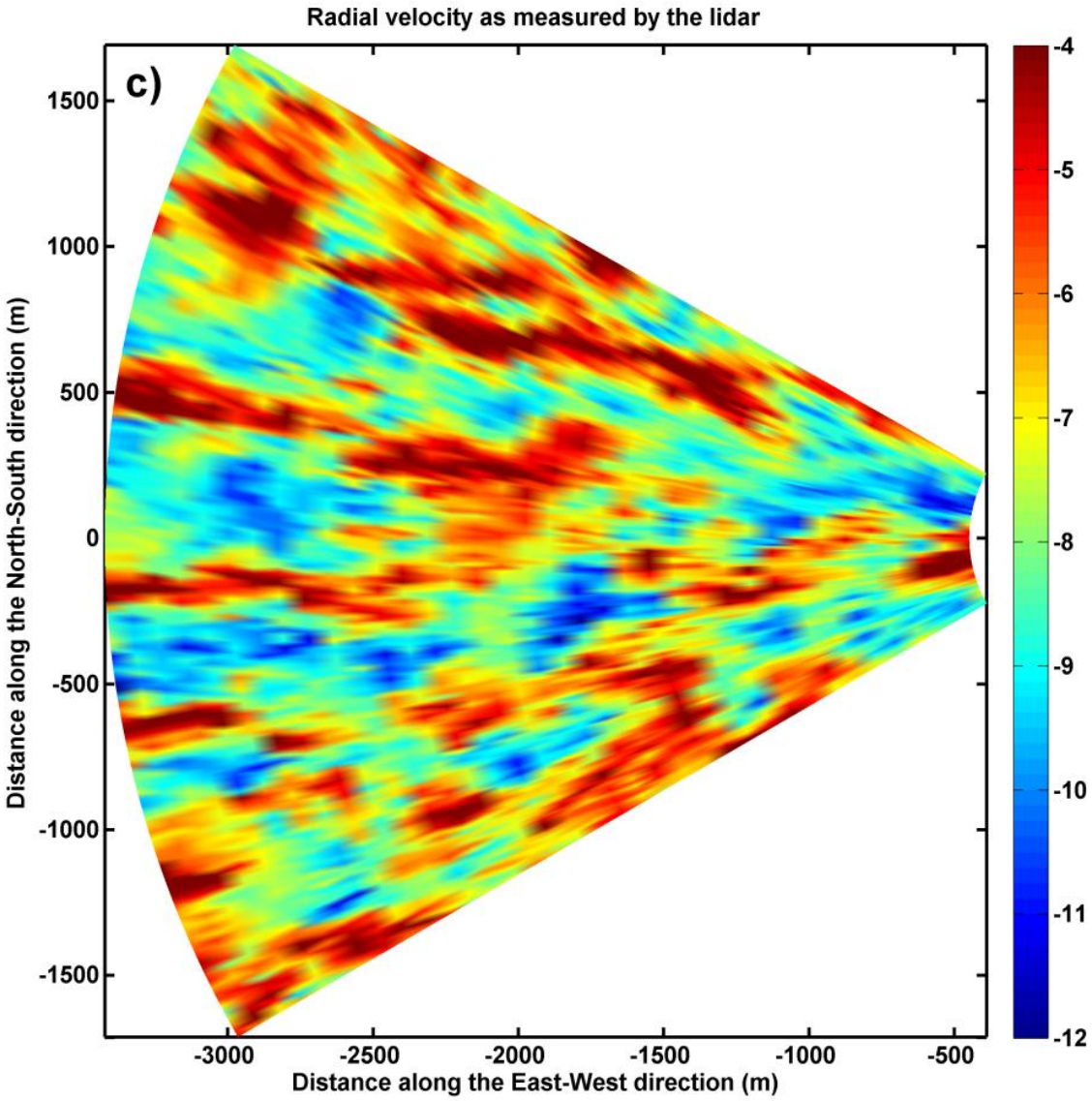
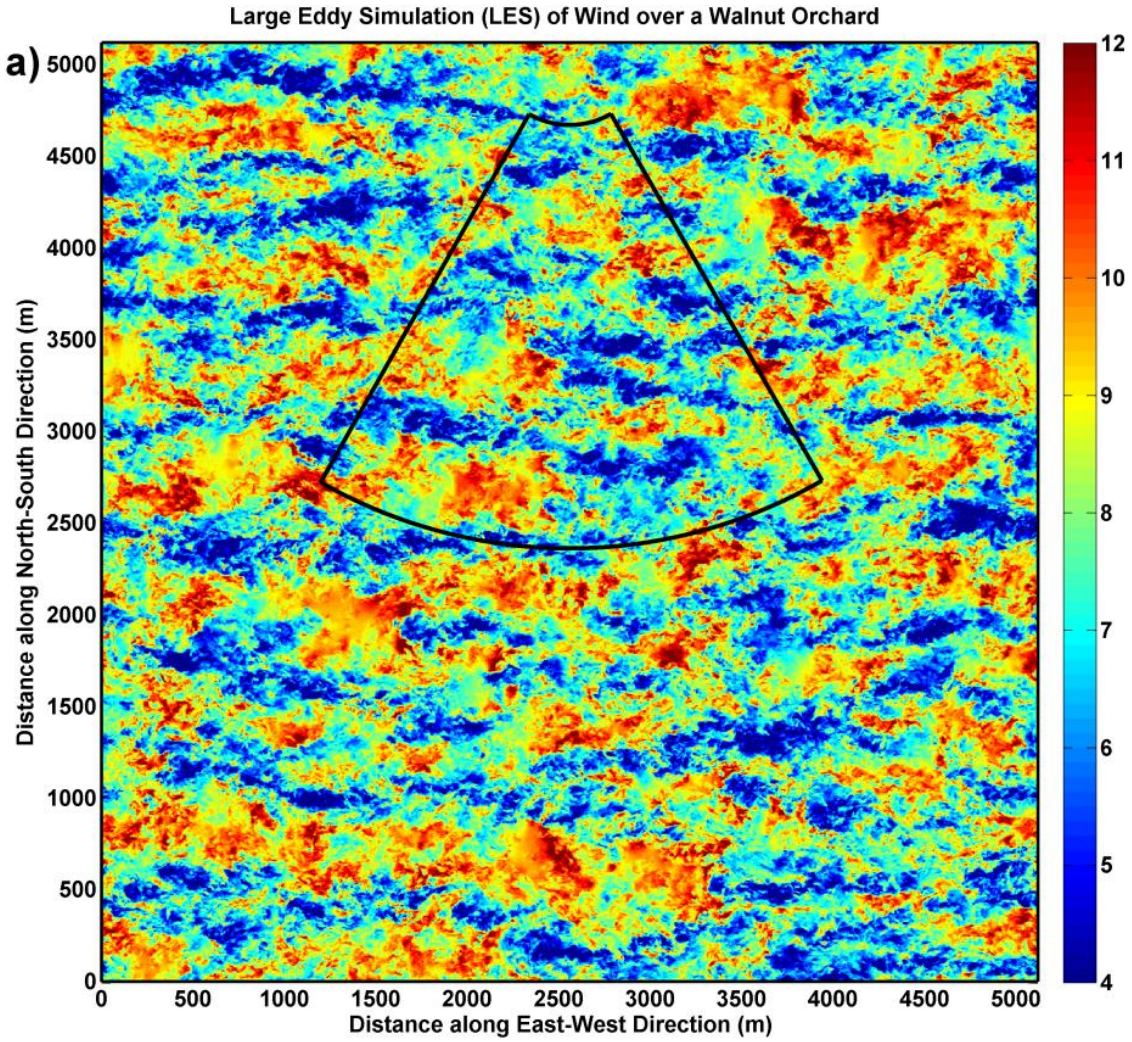
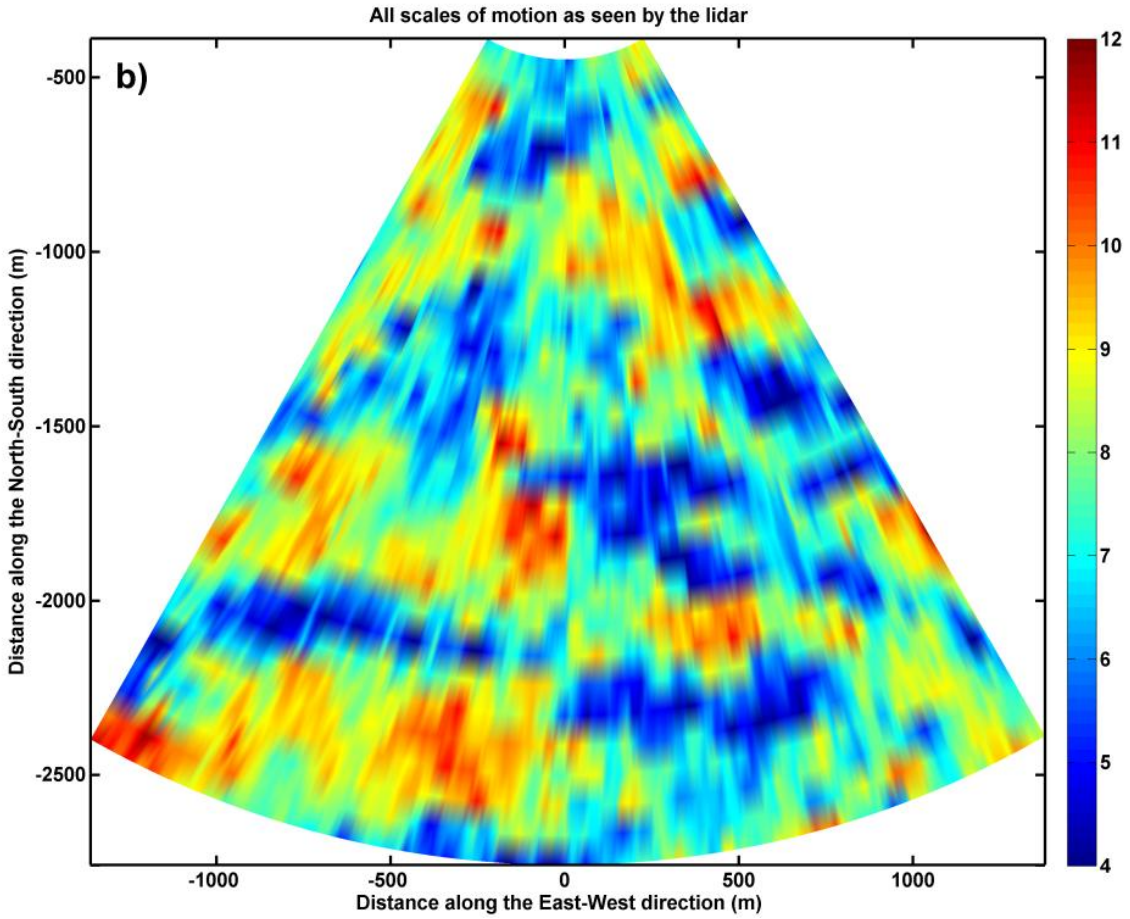


Figure 5. Measurements made using lidar simulator looking in the East-West direction. (a) The actual LES flow field. The black outlines represents the area scanned using the lidar simulator. Colors indicate magnitude of wind speed in m s^{-1}. (b) The LES field as seen by the lidar. Colors indicate magnitude of wind speed in m s^{-1}. (c) Lidar measured radial velocity of the LES field. Colors indicate radial velocity magnitude in m s^{-1}. It

should be noted that the velocity magnitude will appear flipped compared to (b) as the radial velocity towards the lidar is negative in magnitude.

It should be noted that in Figure 5, the LES velocity field evolves as the lidar simulator scans through it. As a result, the velocity structures seen in Figure 5(a) are seen to have advected in the lidar scan in Figure 5(b) as azimuth angles increases. The current update rate for the LES field is once every 5 seconds at present. As a result, the velocity field captured by the lidar simulator will show discontinuities when the LES field "evolves" after every 5 seconds. A new run with a higher update rate of 1 Hz is being created and will be used for future analysis. Figure 5(c) shows the radial velocity measured by the lidar simulator which serves as input to the vector retrieval algorithms. Similarly, the LES field is samples perpendicular to the wind field as at a 45^{o} angle as shown in


Large Eddy Simulation (LES) of Wind over a Walnut Orchard
a)
Distance along North-South Direction (m)
Distance along East-West Direction (m)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
4
5
6
7
8
9
10
11
12


All scales of motion as seen by the lidar
b)
-500
-1000
-1500
-2000
-2500
Distance along the North-South direction (m)
-1000
-500
0
500
1000
Distance along the East-West direction (m)
12
11
10
9
8
7
6
5
4

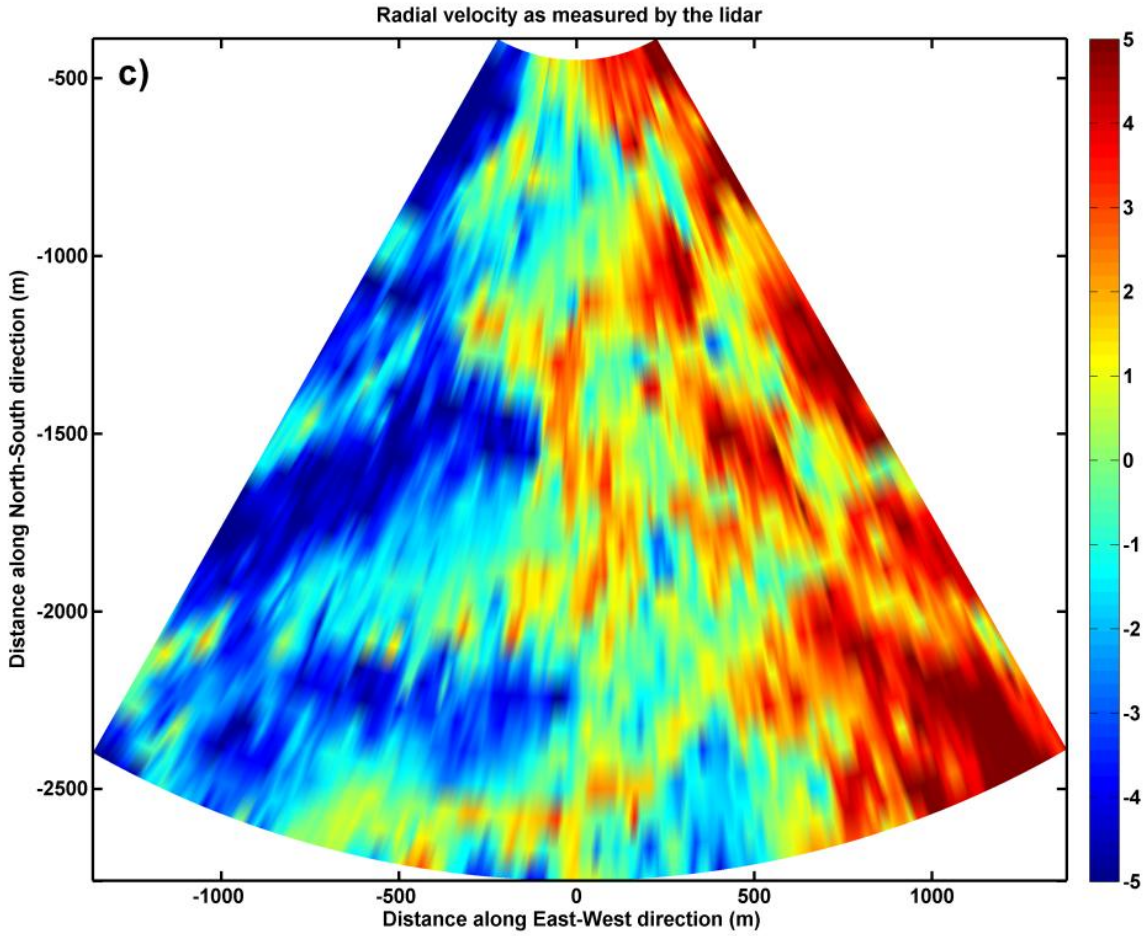
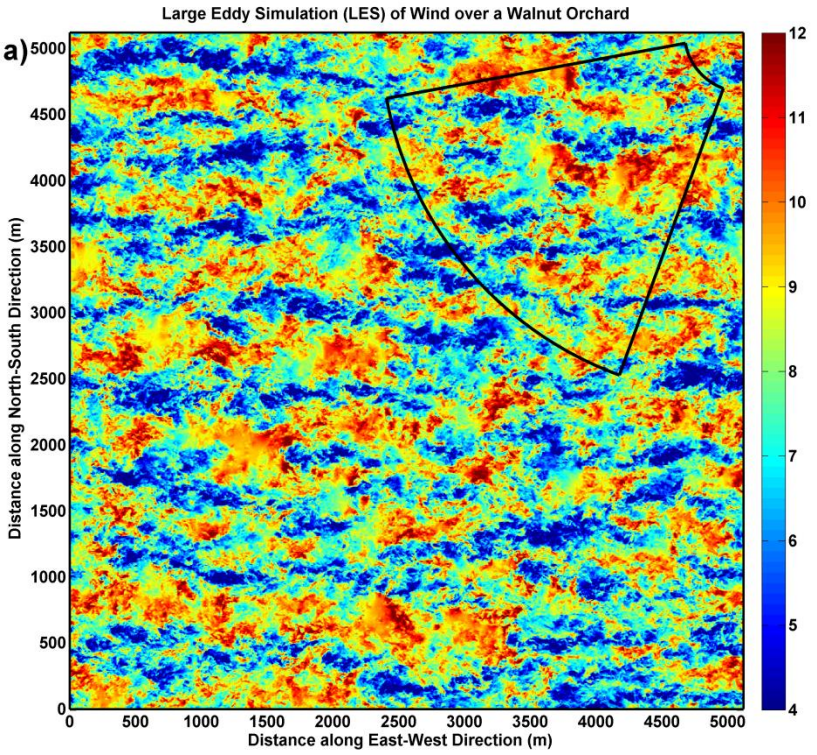
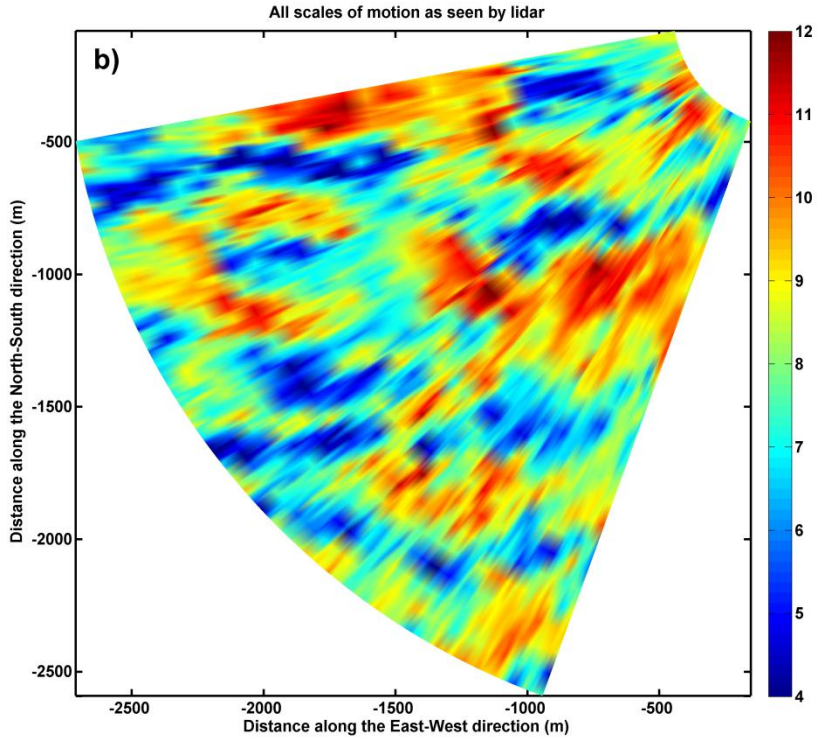


Figure 6. Measurements made using lidar simulator looking in the North-South direction. (a) The actual LES flow field. The black outlines represents the area scanned using the lidar simulator. Colors indicate magnitude of wind speed in m s^{-1}. (b) The LES field as seen by the lidar. Colors indicate magnitude of wind speed in m s^{-1}. (c) Lidar measured radial velocity of the LES field. Colors indicate radial velocity magnitude in m s^{-1}.


Large Eddy Simulation (LES) of Wind over a Walnut Orchard
a)
Distance along North-South Direction (m)
Distance along East-West Direction (m)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
4
5
6
7
8
9
10
11
12


All scales of motion as seen by lidar
b)
Distance along the North-South direction (m)
Distance along the East-West direction (m)
-500
-1000
-1500
-2000
-2500
4
5
6
7
8
9
10
11
12

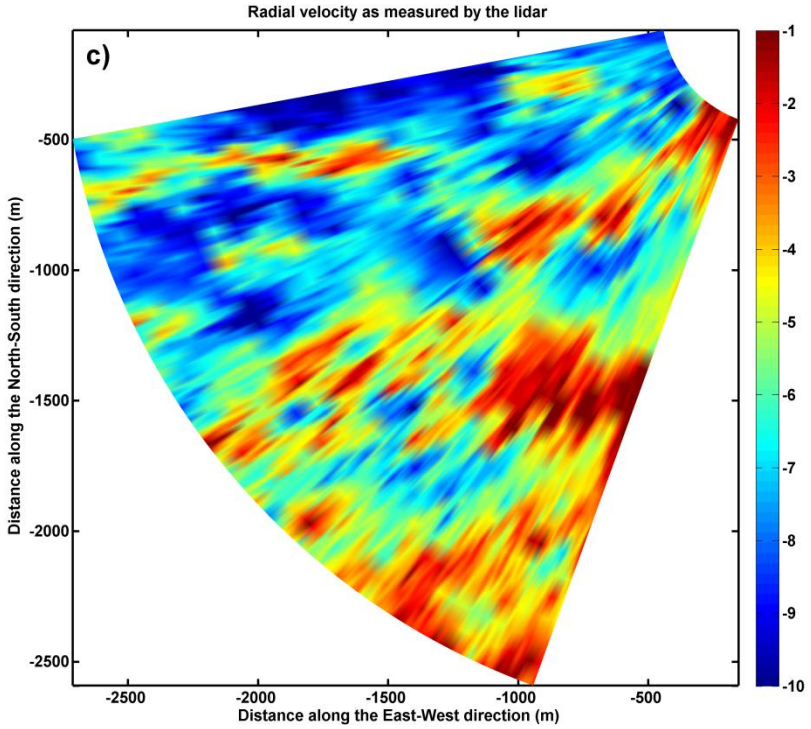
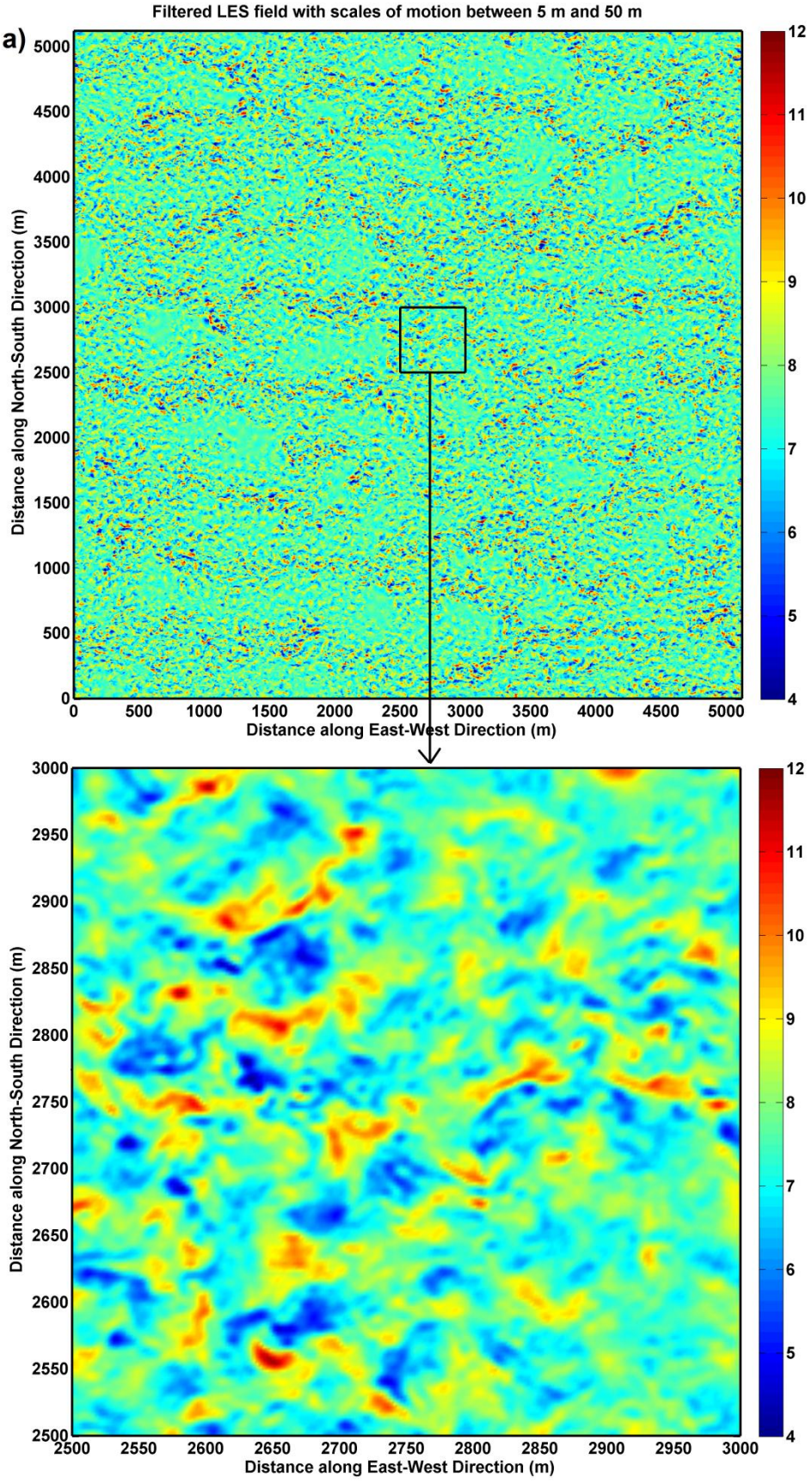


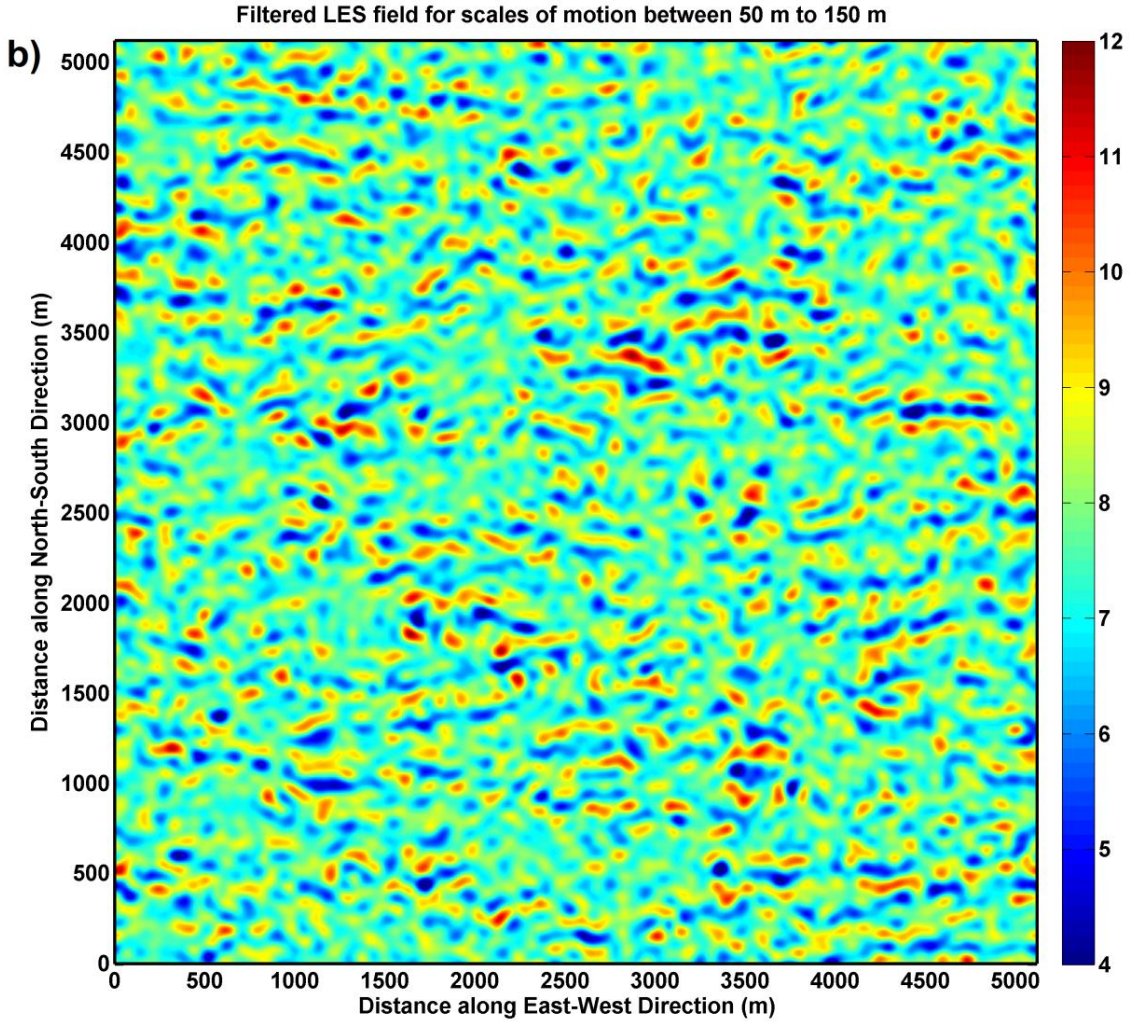
Figure 7. Measurements made using lidar simulator looking from the North-East direction. (a) The actual LES flow field. The black outlines represent the area scanned using the lidar simulator. Colors indicate magnitude of wind speed in m s^{-1}. (b) The LES field as seen by the lidar. Colors indicate magnitude of wind speed in m s^{-1}. (c) Lidar measured radial velocity of the LES field. Colors indicate radial velocity magnitude in m s^{-1}.

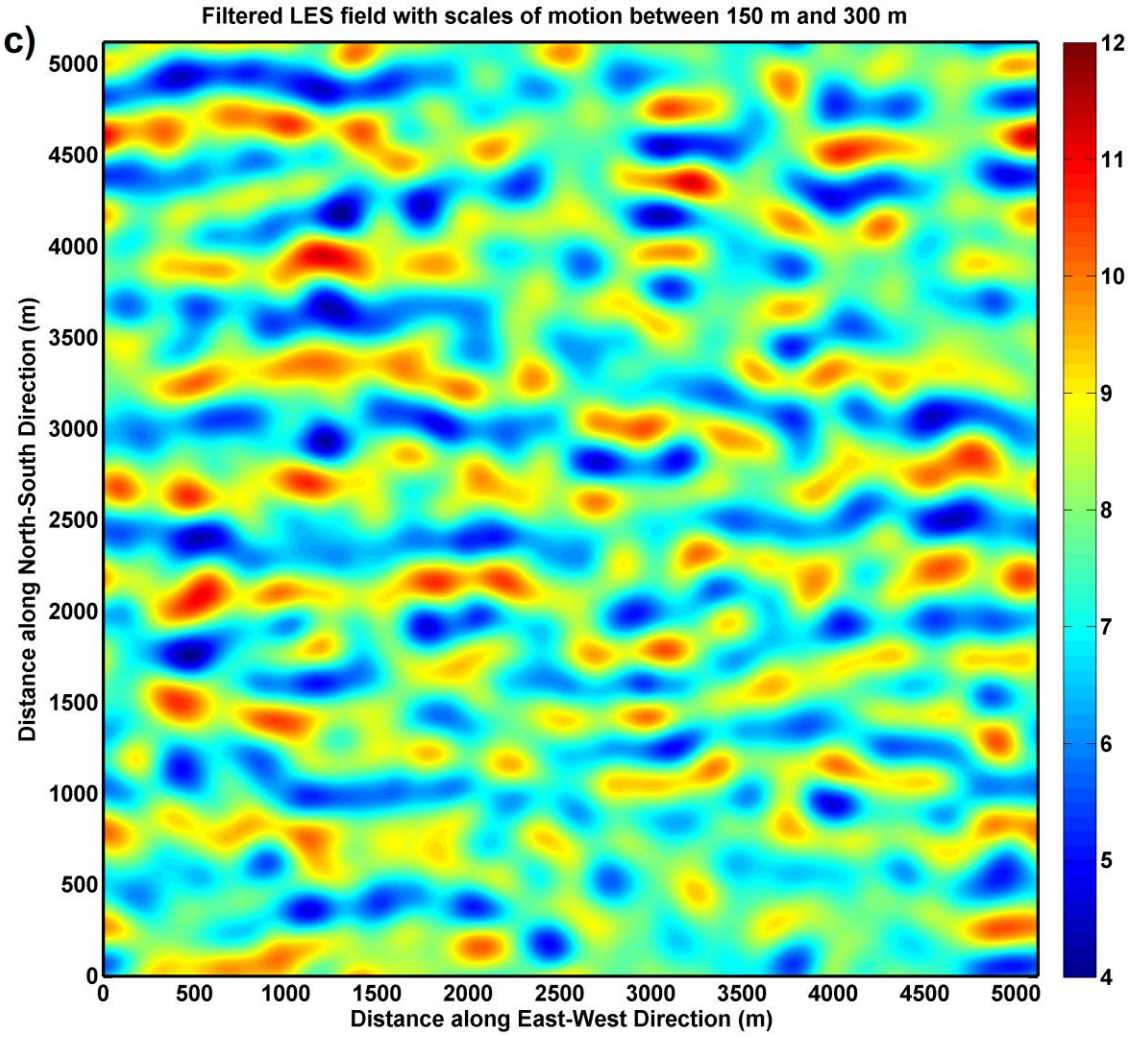
4.3.2 Effect of scales of motion

In order to quantify the error of representativeness of the lidar system, the LES field is frequency filtered to contain scales of motion in the following ranges: (1) scales between 5 m to 50 m, (2) scales between 50 m and 150 m, (3) scales between 150 m and 300 m and (4) scales above 300 m. The lidar simulator was setup to have a range-gate size of 80 m, which is a typical value on many ASU lidar deployments using the Wind

Tracer. With this setup, the lidar will either not capture the scales of motion between 5 m and 50 m, or misrepresent them. The scales of motion between 50 m and 150 m are expected to be partially captured, while scales of motion above 150 m should be well captured. This hypothesis can be tested by filtering the LES data as explained above and quantifying the error in lidar measurement in each case. The LES data filtered based on the above criteria are shown in Figure 8.


Filtered LES field with scales of motion between 5 m and 50 m
a)
Distance along North-South Direction (m)
Distance along East-West Direction (m)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
4
5
6
7
8
9
10
11
12
2500
2550
2600
2650
2700
2750
2800
2850
2900
2950
3000


Filtered LES field for scales of motion between 50 m to 150 m
b)
Distance along North-South Direction (m)
Distance along East-West Direction (m)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
4
5
6
7
8
9
10
11
12


c)
Filtered LES field with scales of motion between 150 m and 300 m
Distance along North-South Direction (m)
Distance along East-West Direction (m)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
4
5
6
7
8
9
10
11
12

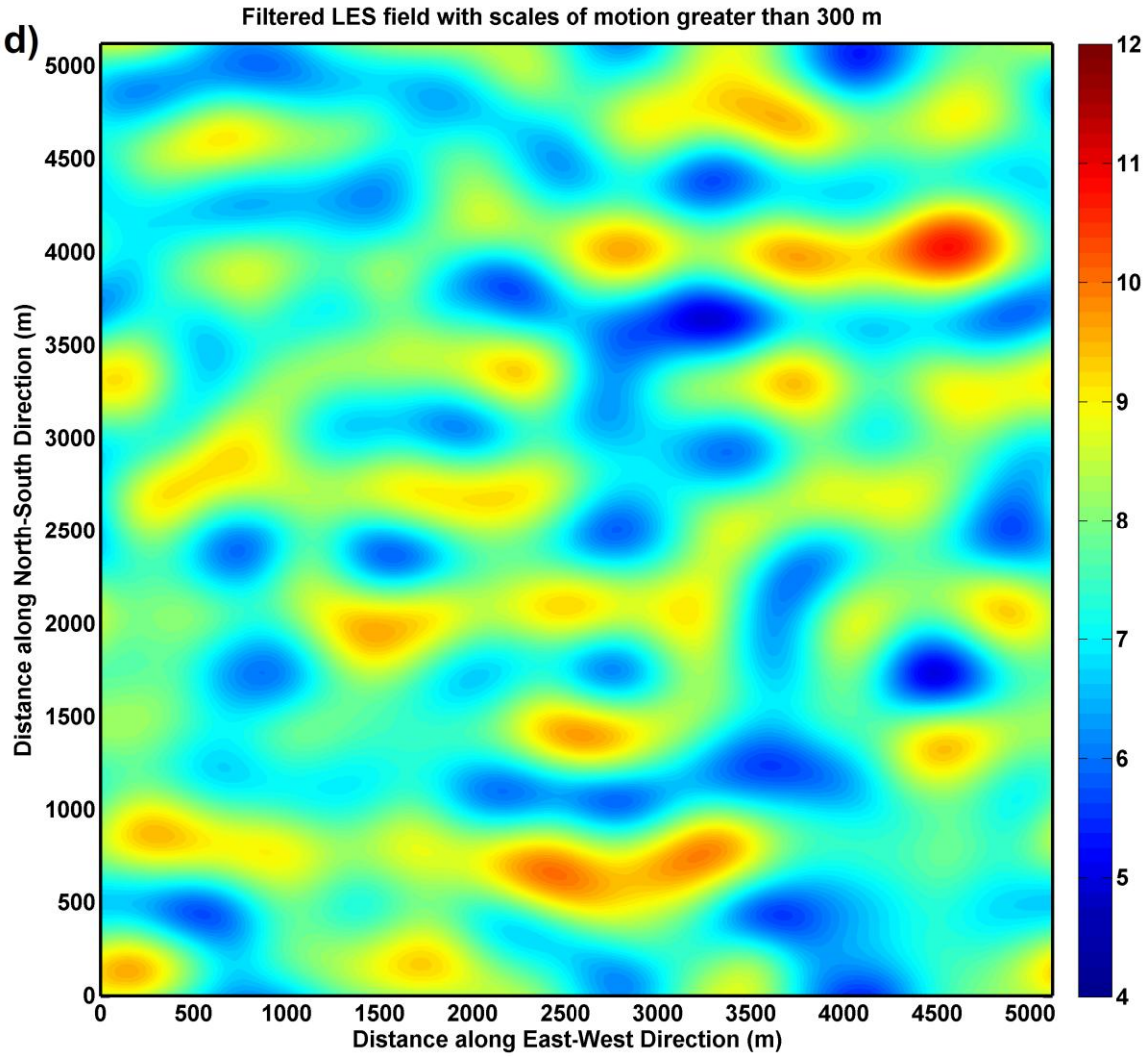
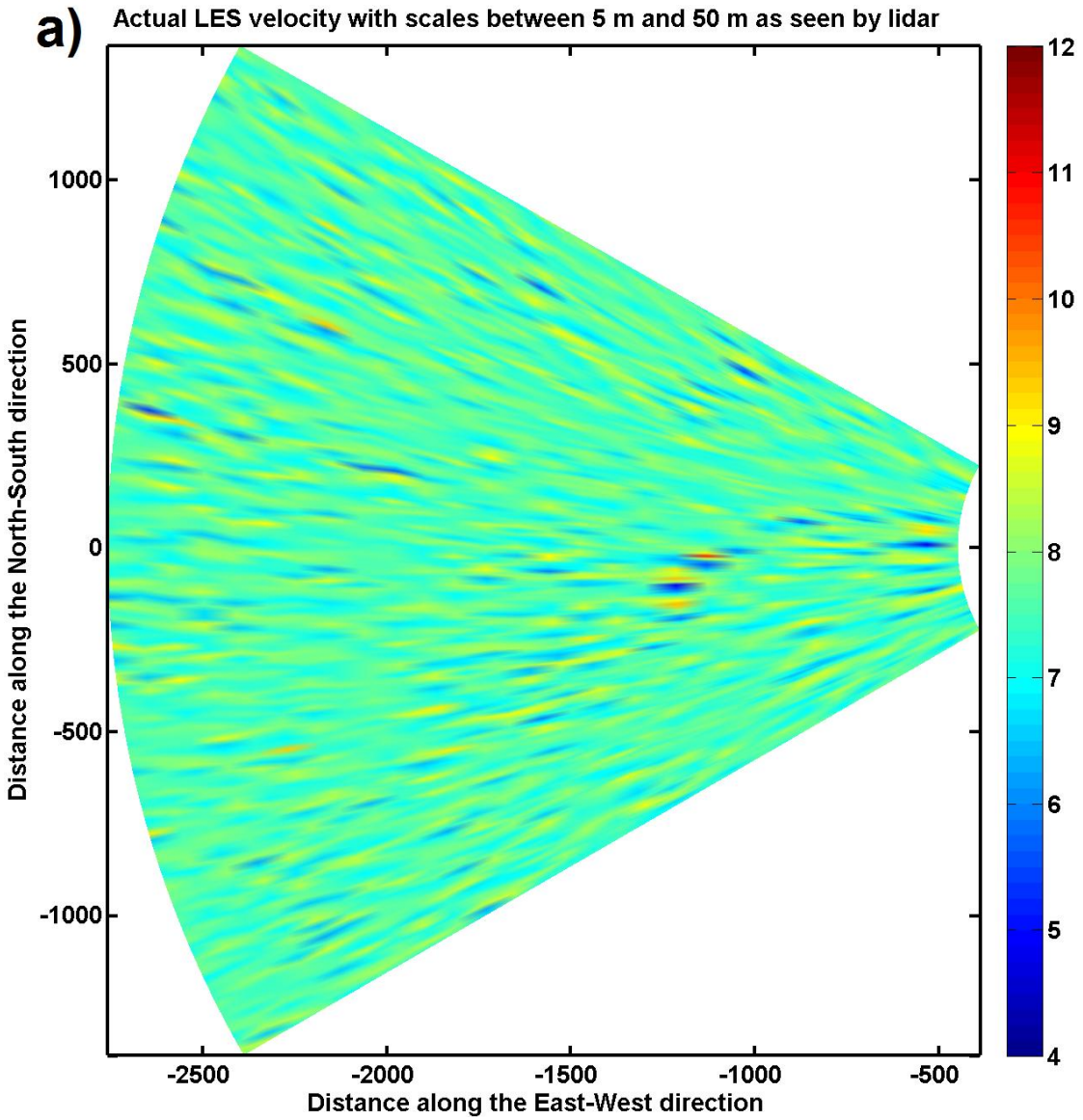


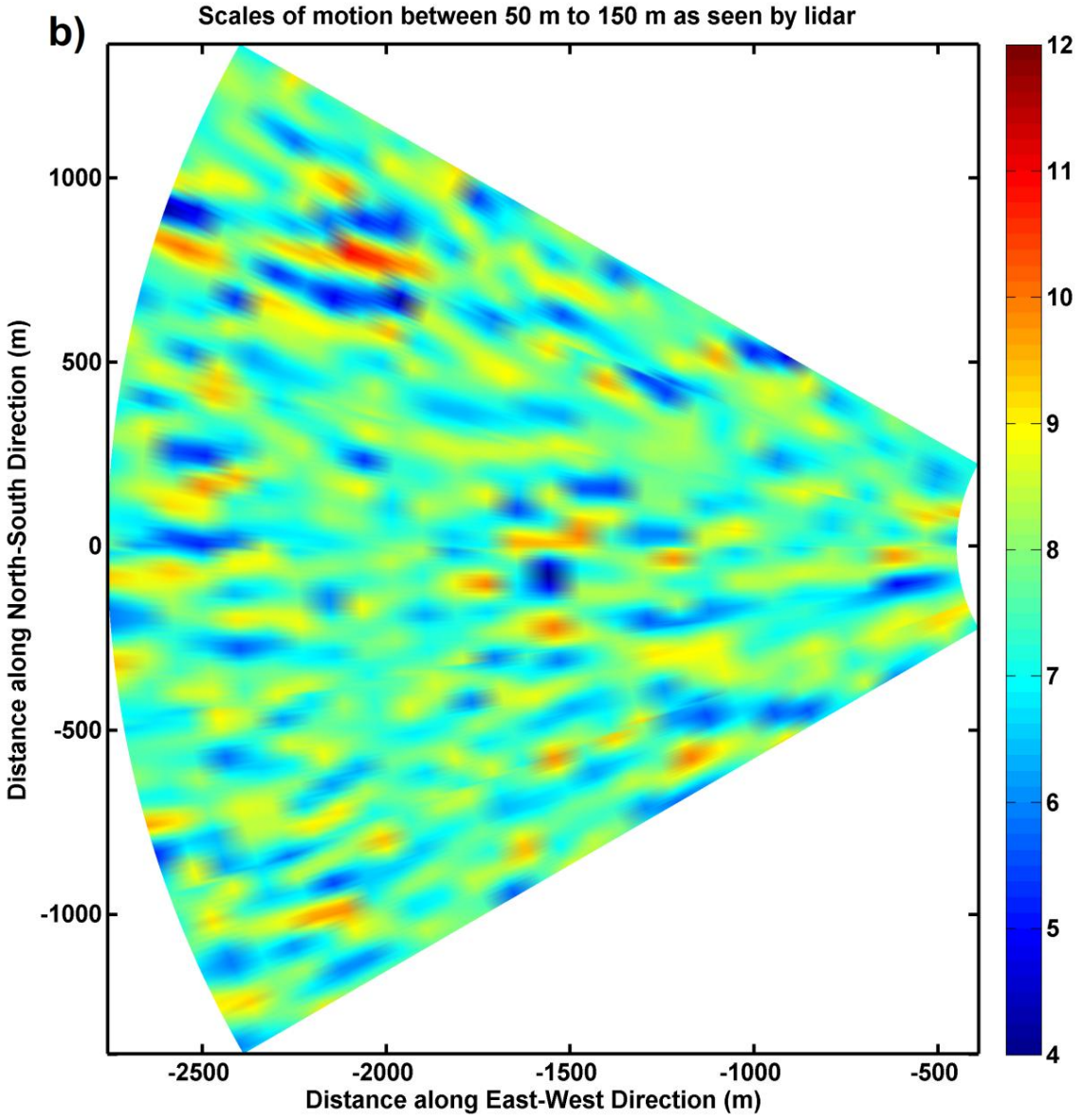
Figure 8. The LES velocity field, filtered to retain different scales of motion. (a) Scales of motion between 5 m and 50 m. (b) Scales of motion between 50 m and 150 m. (c) Scales of motion between 150 m and 300 m. (d) Scales of motion above 300 m. The colors indicate the magnitude of wind speed.

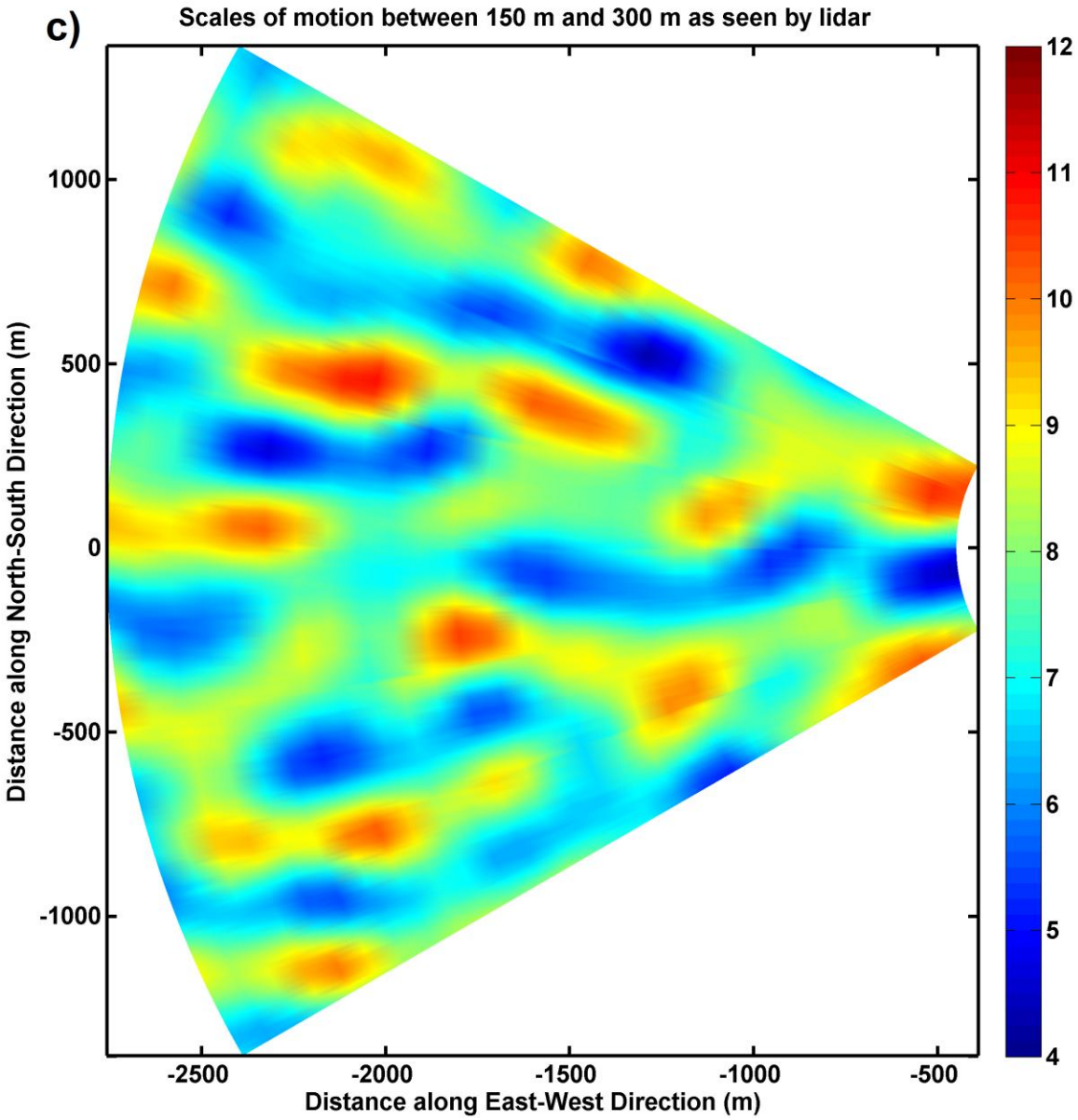
4.4 Error of Representativeness

Our observation of the universe is limited to the range of scales resolved by the measurement platform. Scales outside of this range are either lost or misrepresented.

Quantifying this error is very difficult with real atmospheric measurements. However, with the lidar simulator, this error can be quantified for different scanning configurations. The lidar simulator scanned the velocity fields with different scales of motion present and the results are presented in Figure 9.




b)
Scales of motion between 50 m to 150 m as seen by lidar
Distance along North-South Direction (m)
1000
500
0
-500
-1000
-2500
-2000
-1500
-1000
-500
Distance along East-West Direction (m)
12
11
10
9
8
7
6
5
4


c)
Scales of motion between 150 m and 300 m as seen by lidar
Distance along North-South Direction (m)
1000
500
0
-500
-1000
-2500
-2000
-1500
-1000
-500
Distance along East-West Direction (m)
12
11
10
9
8
7
6
5
4

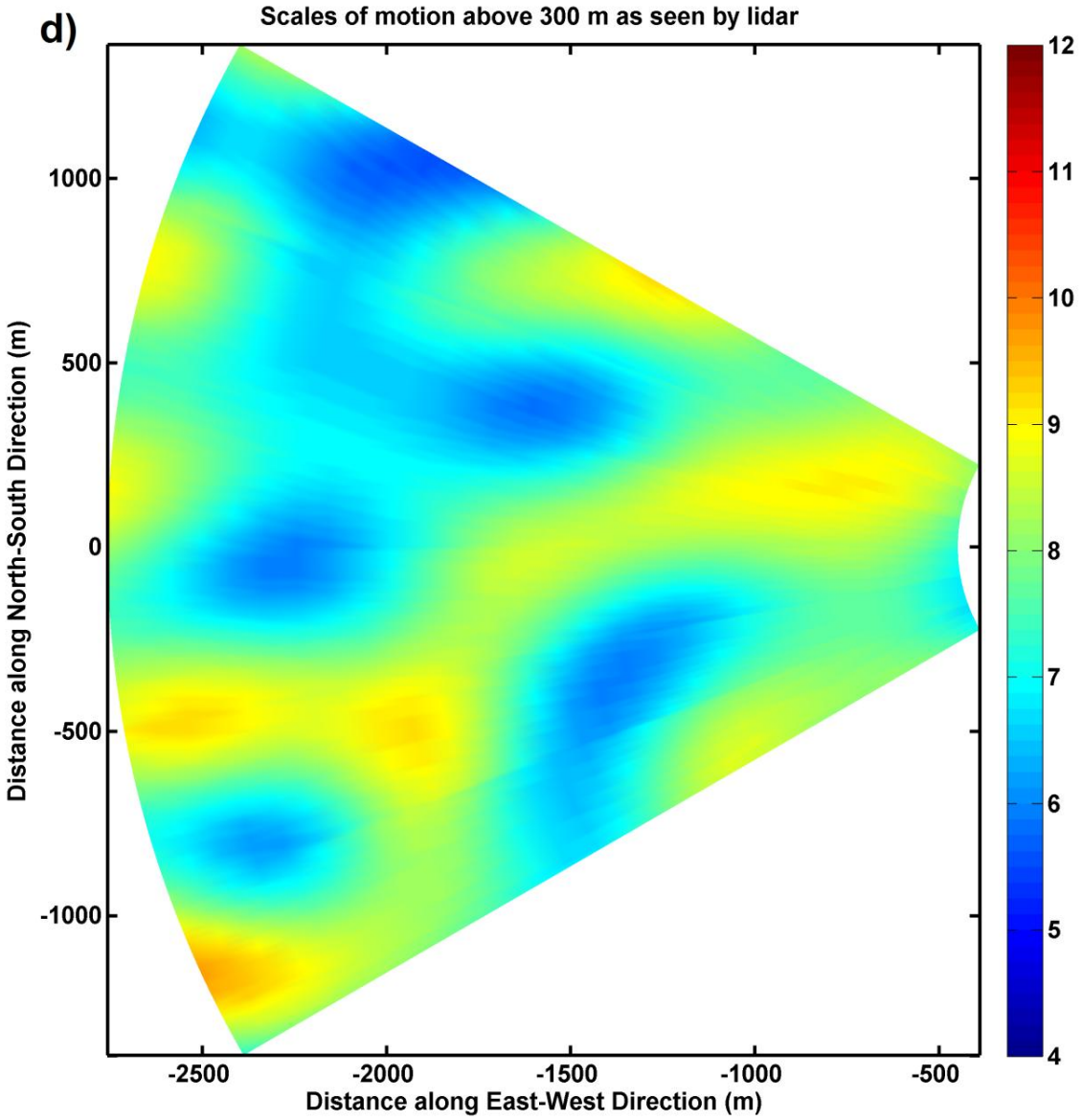


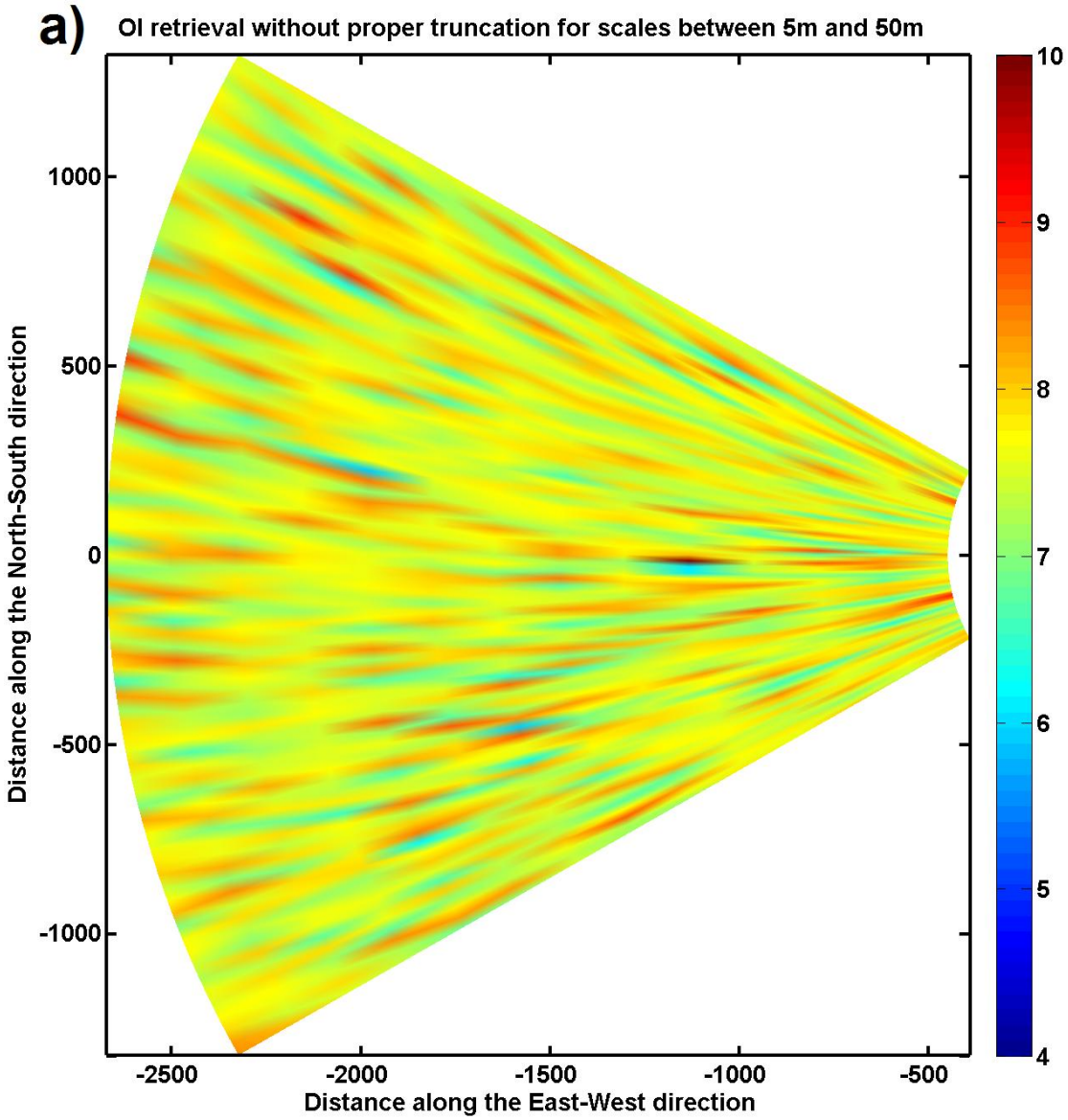
Figure 9. Results from lidar simulator after scanning wind fields with different scales of motion. (a) Scales of motion between 5 m and 50 m. (b) Scales of motion between 50 m and 150 m. (c) Scales of motion between 150 m and 300 m. (d) Scales of motion above 300 m. Colors indicate magnitude of wind speed.

Figure 9(a) shows the lidar simulator results for the scales of motion between 5 m and 50 m. These scales of motion are less than the size of the range-gate and hence will

be under-sampled. As a result, they form the artificial streaks that are seen in Figure 9(a). Figure 9(b) shows lidar simulator results for scales of motion between 50 m and 150 m. These scales are at size just below the range-gate size or greater than range-gate size and hence are better resolved than the scales in Figure 9(a). Although there is still some misrepresentation seen in the how the scales were captured by the lidar simulator. Finally in Figure 9(c) and Figure 9(d), show the lidar simulator results for scales of motion between 150 and 300 m and greater than 300 m respectively. It can be seen from these two figures that these scales are captured quite accurately by the lidar as they are significantly larger than the range-gate size.

4.5 Vector Retrieval

The results from the lidar simulator are analyzed using the modified optimal interpolation (OI) technique. The velocity vectors retrieved from the OI technique are compared to the velocity vectors as seen by the lidar using the lidar simulator to determine the velocity retrieval error. This error is expected to be a function of the size of scales and the look angle. As observed in Figure 9(a), the lidar simulator produces unnatural streaks as a result of error of representativeness for scales between 5 m and 50 m. When performing vector retrieval on such a product, it is desirable that the vector retrieval algorithm filter out these artifacts of the lidar measurement process. With the OI technique, proper selection of the truncation number is seen to filter out these artifacts. The retrievals using the OI technique and the effect of truncation number for scales of motion between 5 m and 50 m is shown in Figure 10.


a)
OI retrieval without proper truncation for scales between 5m and 50m
Distance along the North-South direction
1000
500
0
-500
-1000
-2500
-2000
-1500
-1000
-500
Distance along the East-West direction
10
9
8
7
6
5
4

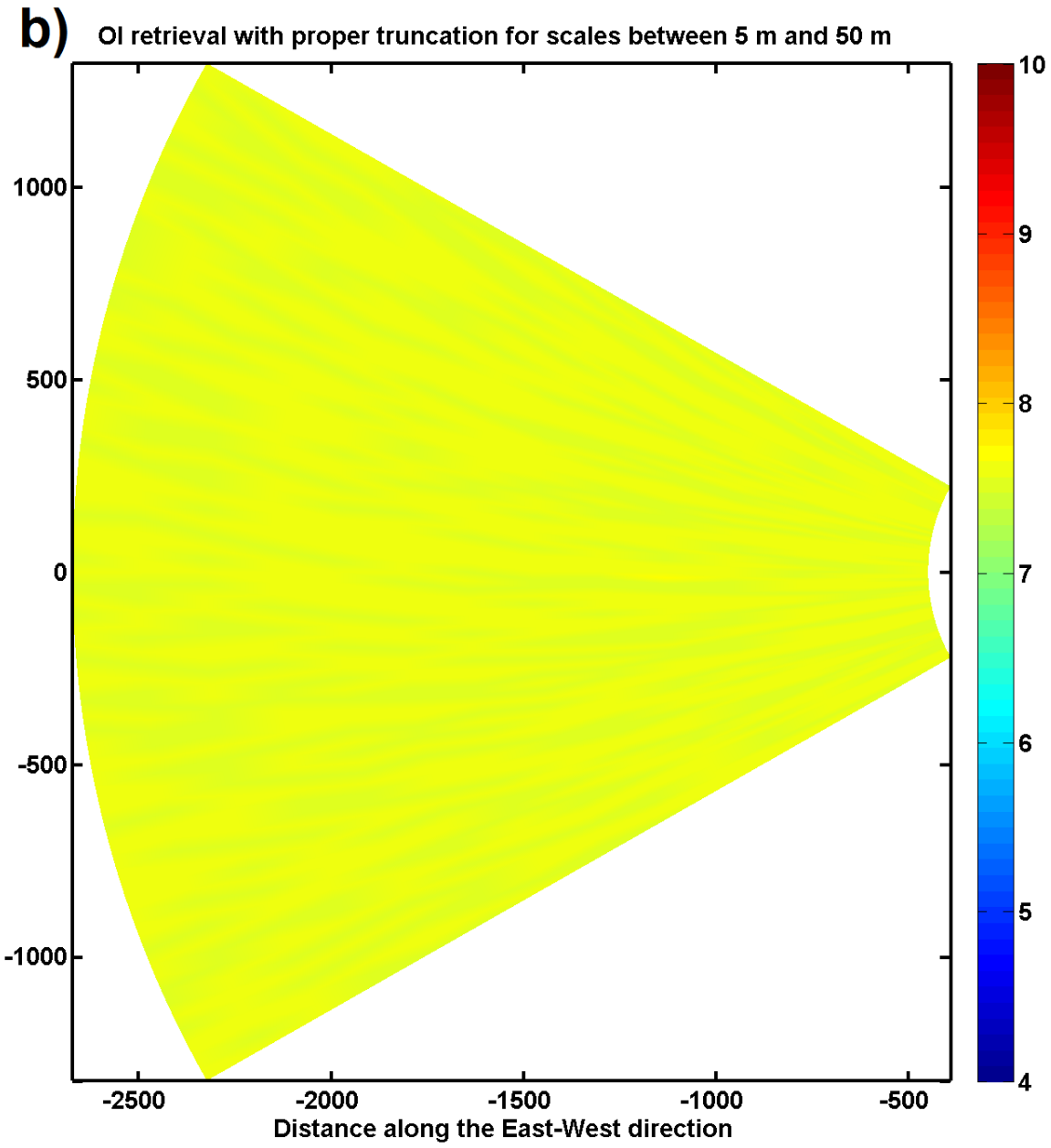
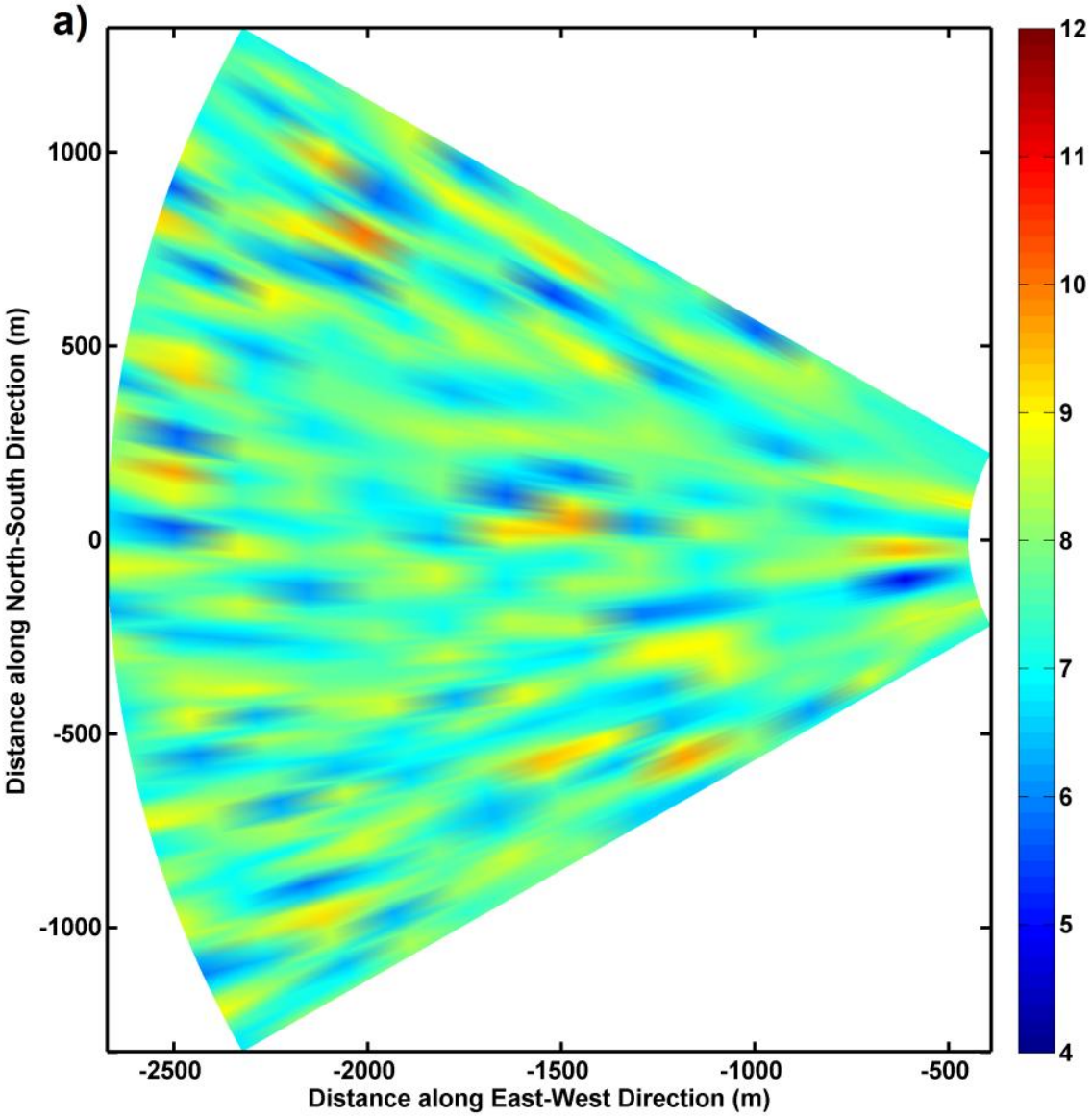
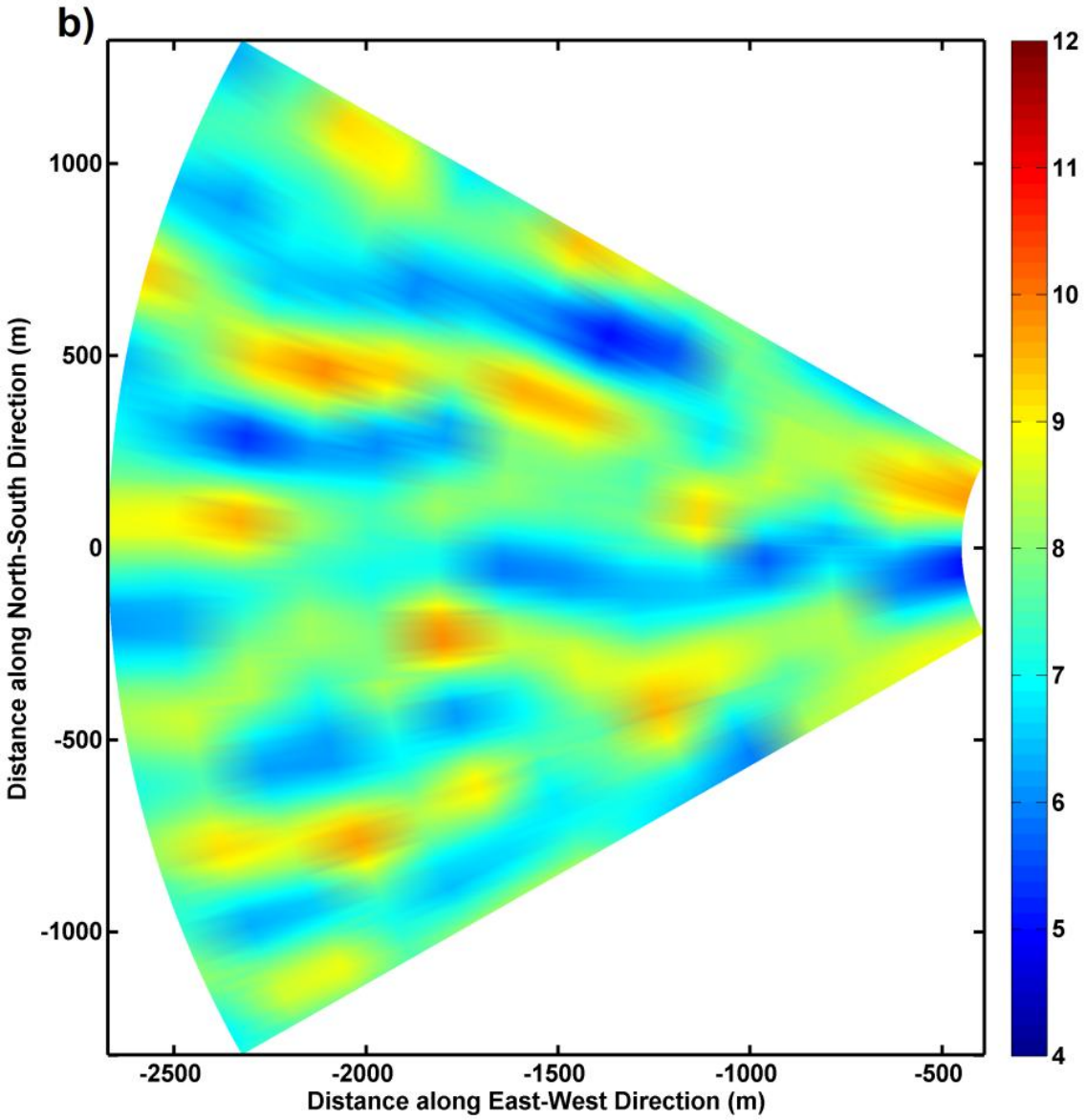


Figure 10. Vector retrieval using the OI technique for scales of motion between 5 m and 50 m. (a) Retrieval using improper truncation number. (b) Retrieval using proper truncation number. The colors indicate magnitude of wind speed.

The OI retrievals for the other scales of motion are shown in Figure 11. These retrievals were performed using the proper truncation number choice as explained above.

As seen from Figure 11, the retrieval process is able to capture a wide range of scales quite accurately. The optimal truncation number is determined through an iterative procedure at the beginning of the analysis and remains constant after that.




b)
1000
500
0
-500
-1000
Distance along North-South Direction (m)
-2500
-2000
-1500
-1000
-500
Distance along East-West Direction (m)
12
11
10
9
8
7
6
5
4

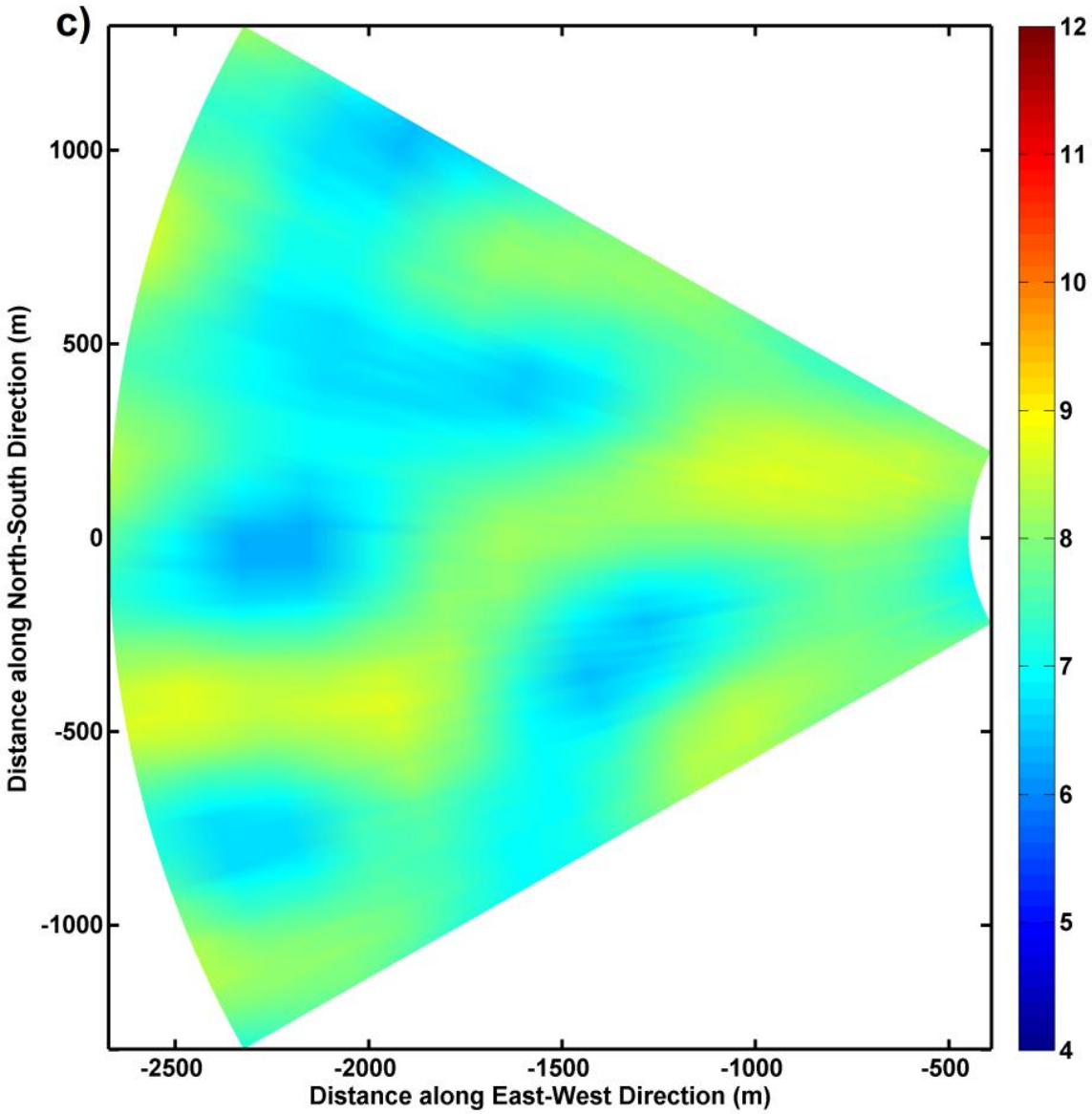
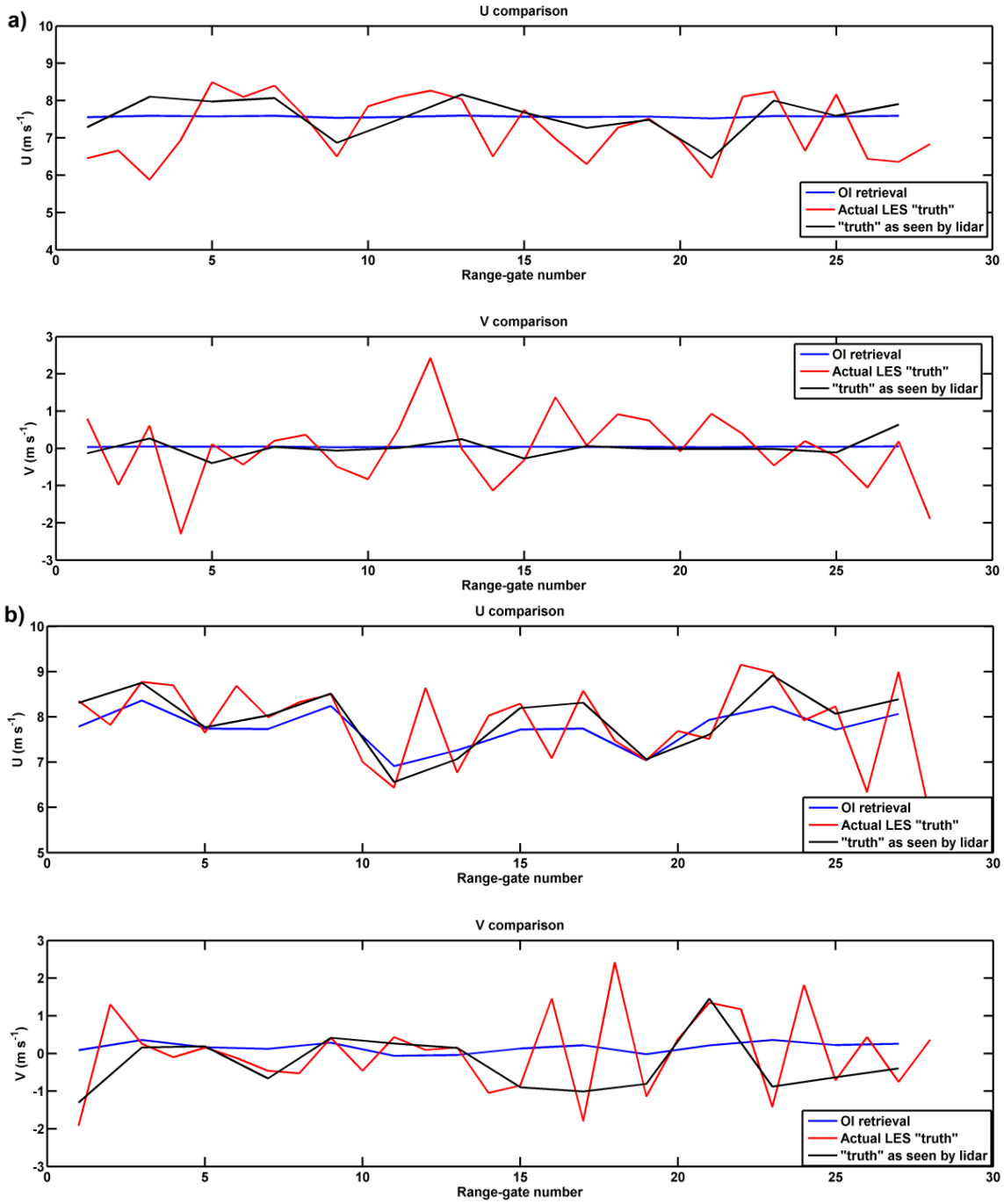


Figure 11. Vector retrieval using the OI technique. (a) Retrieval for scales between 50 m and 150 m. (b) Retrieval for scales between 150 m and 300 m. (c) Retrieval for scales above 300 m. Colors indicate magnitude of wind speed.

4.6 Understanding Error of Representativeness and Retrieval Error

In order to distinguish between the lidar error of representativeness and vector retrieval error, the time series of the "truth" velocity vectors is compared to the lidar

sampled velocity and the OI retrieved velocity as shown in Figure 12. As seen from Figure 12, the error of representativeness (difference between the red-line and the black-line) is highest for scales of motion between 5 m and 50 m. It reduces for scales between 50 m and 150 m and even more for scales between 150 m and 300 m. It is the lowest for scales above 300 m.


a)
U comparison
U (m s-1)
Range-gate number
OI retrieval
Actual LES "truth"
"truth" as seen by lidar
V comparison
V (m s-1)
Range-gate number
OI retrieval
Actual LES "truth"
"truth" as seen by lidar
b)
U comparison
U (m s-1)
Range-gate number
OI retrieval
Actual LES "truth"
"truth" as seen by lidar
V comparison
V (m s-1)
Range-gate number
OI retrieval
Actual LES "truth"
"truth" as seen by lidar

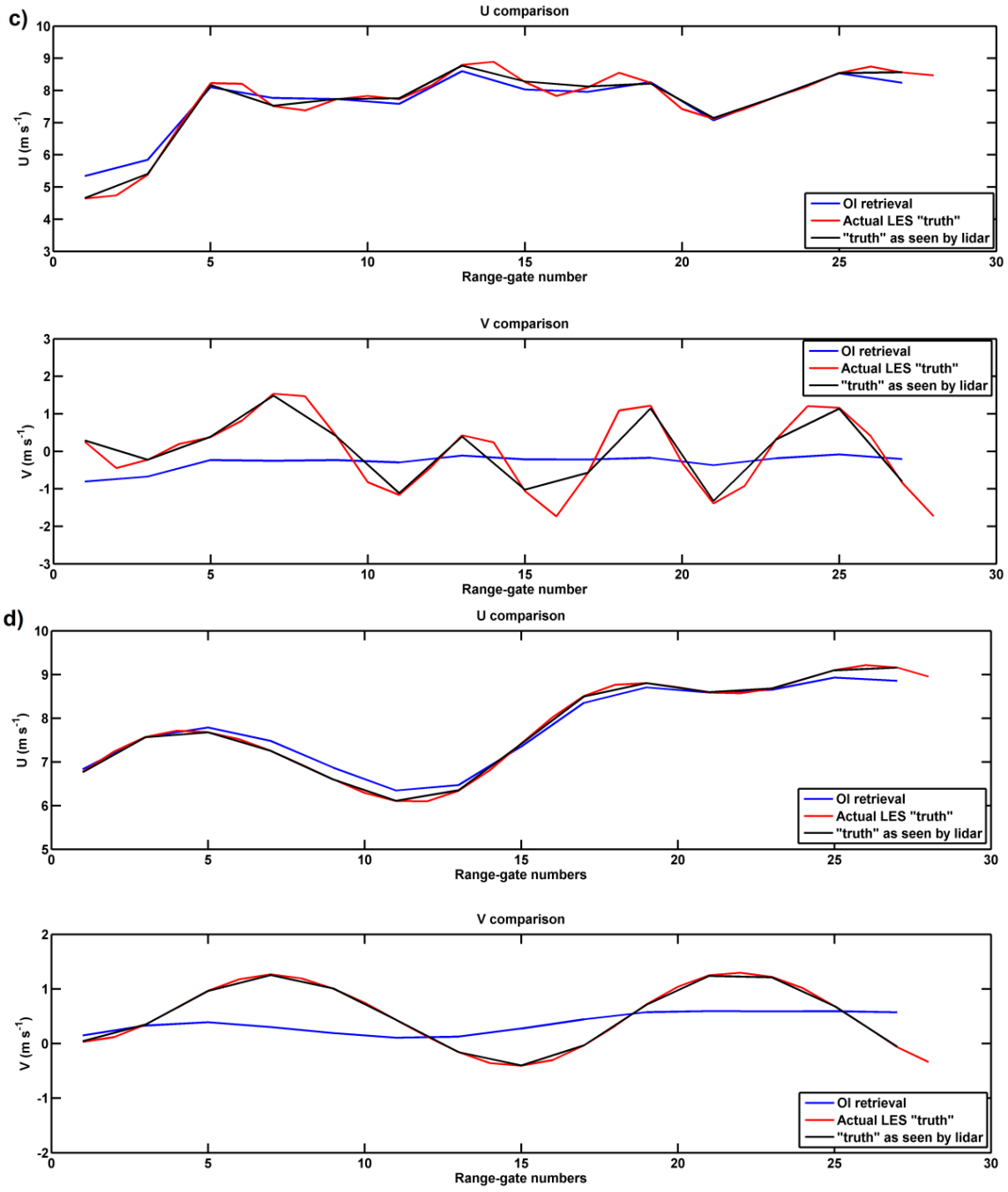


Figure 12. Comparison of lidar sampled velocities and the OI retrieved velocities to the LES "truth" velocities. (a) Comparison for scales of motion between 5 m and 50 m. (b) Comparison for scales of motion between 50 m and 150 m. (c) Comparison for scales of motion between 150 m and 300 m. (d) Comparison for scales of motion above 300 m.

In this case, the lidar simulator was setup to have a range-gate size of 80 m which is consistent with the range-gate size setup during the Canopy Horizontal Array Turbulence Study (CHATS) campaign. As a result, scales of motion that are smaller than the range-gate size are misrepresented due to insufficient sampling. Therefore, when making atmospheric observations, measurements of the resolvable scales of motion are contaminated by error of representativeness of the unresolved scales. Hence, appropriate post-processing techniques need to be applied that can account for this misrepresentation.

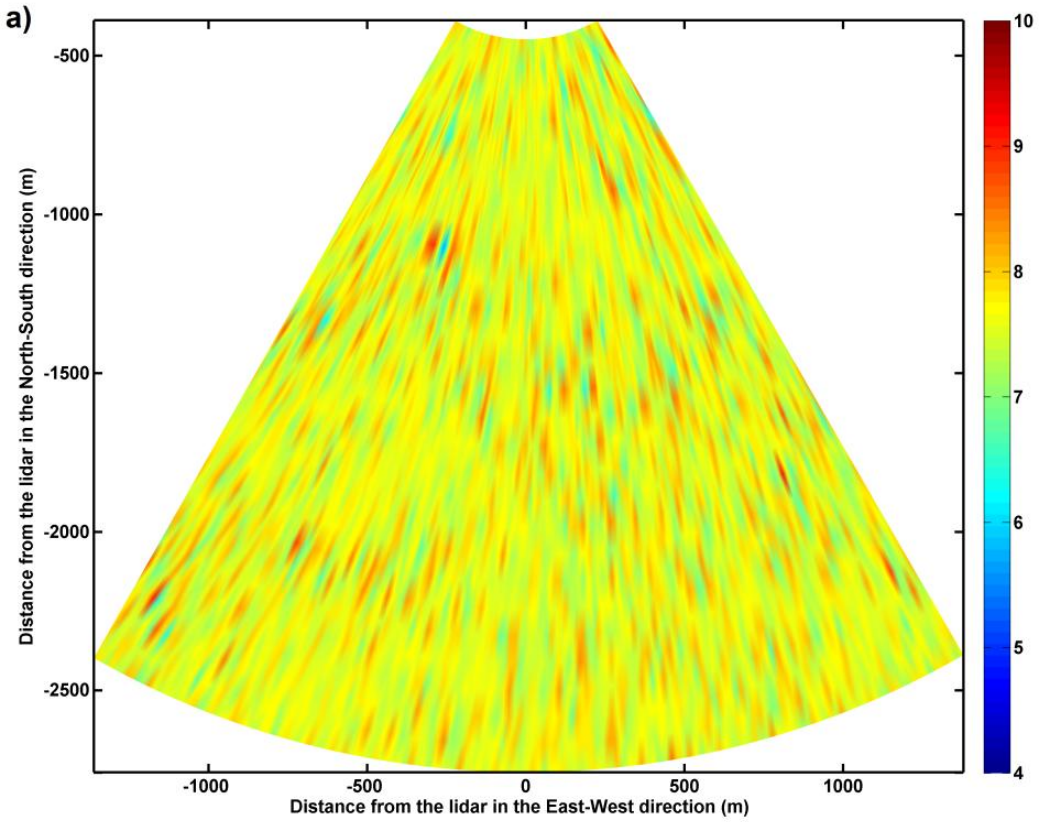
Using the modified optimal interpolation technique, this error of representativeness can be accounted for through proper selection of truncation number (see Section 2.4.2). The effect of proper selection of truncation number is demonstrated in Figure 10. In Figure 12(a), the results from the modified OI technique are seen to filter out the lidar measurements artifacts and retrieve the mean wind speed.

4.7 Effect of look angle

The coherent Doppler lidar measures only the radial component of the wind speed. The magnitude of the component captured depends on the angle of the beam with respect to the wind direction and approaches zero as the wind direction becomes normal to the beam direction. As a result, vector retrieval becomes more challenging when the direction of the wind is perpendicular to beam direction. In order to investigate the sensitivity of vector retrieval to look angle, lidar simulator was used to simulate measurements from different look angles. Using this analysis, the vector retrieval accuracy will be quantified as a function of look angle.

To estimate the retrieval accuracy as function of look angle, the lidar simulator was first setup to measure the flow with a look angle close to perpendicular to the flow. As in the previous cases, the velocity field was filtered to have different scales of motion present. Therefore the effect of look angle on retrieval of different scales of motion can be estimated. Figure 13 shows the lidar measurement and vector retrieval of the flow with scales of motion between 5 m and 50 m and look angle close to perpendicular. A misrepresentation of the flow is observed as in the previous cases as the scales of motion are smaller than the range-gate size. The modified OI technique is able to filter out the artifacts due to error of representativeness as explained previously. This result is similar to the one retrieved when the lidar was looking into the flow as both retrieve only the mean wind.

Figure 14(a) shows the scales of motion between 50 m and 150 m as seen by the lidar. Since some of the scales of motion are below the range-gate size, there is misrepresentation of the scales of motion. Comparing Figure 9(b) and Figure 14(a) we see that the artifacts due to error of representativeness also depend on the look angle. However, as seen in Figure 14(b), the modified OI technique is able to filter out most of the artifacts.


a)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

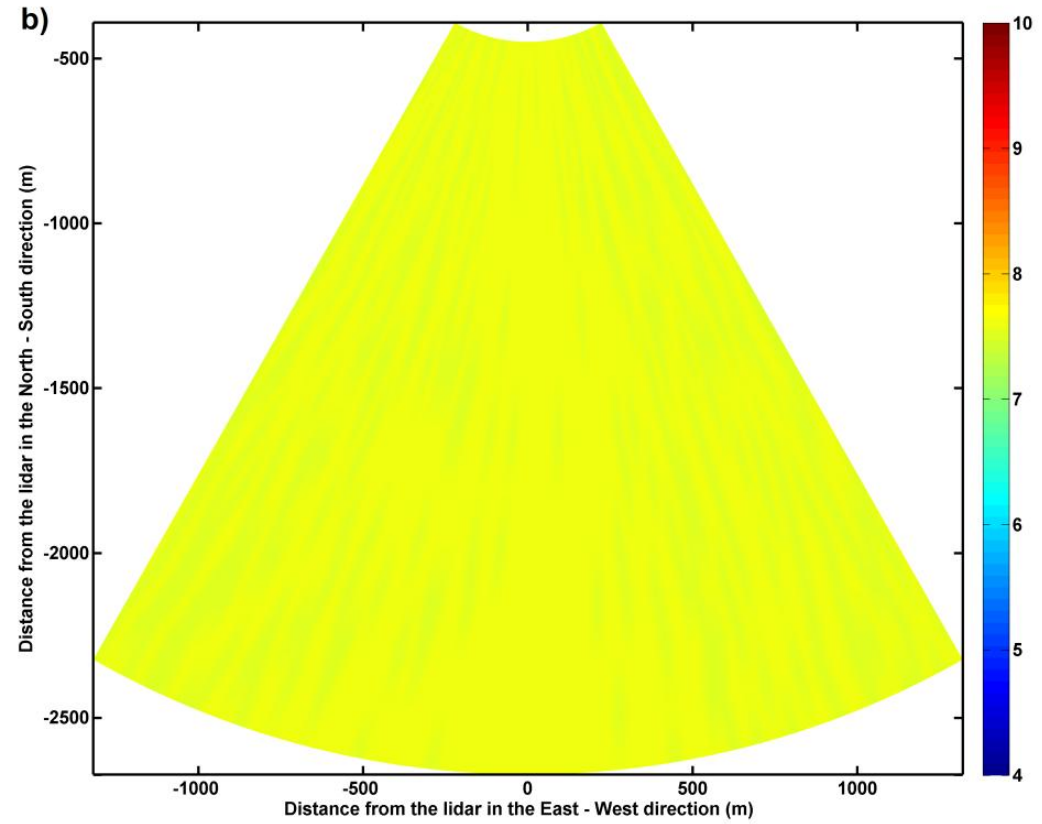

b)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North - South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East - West direction (m)
10
9
8
7
6
5
4

Figure 13. Lidar measurement of the flow with scales of motion between 5 m and 50 m and look angle close to perpendicular. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion between 5 m and 50 m with proper truncation number. The colors indicate magnitude of wind speed.

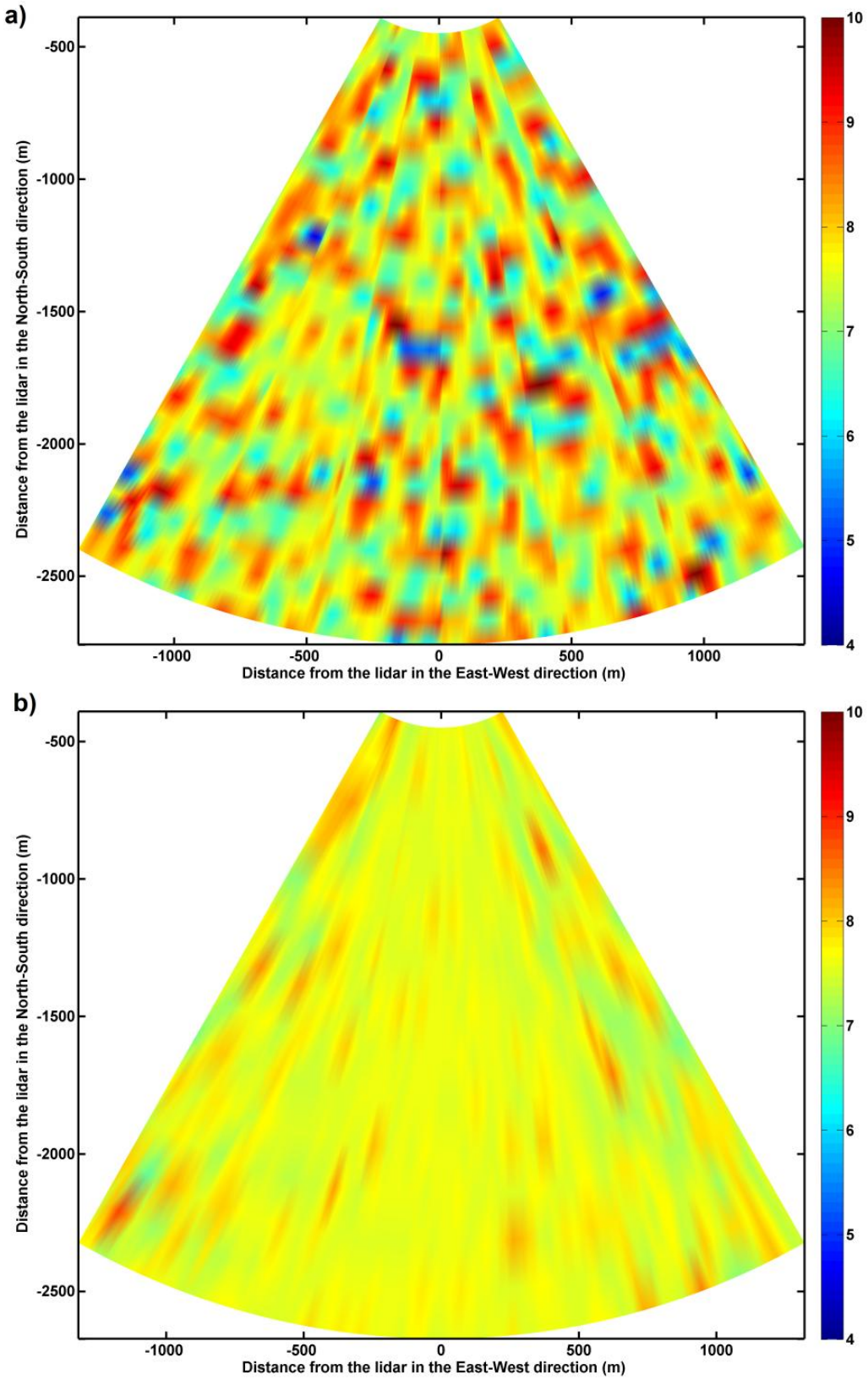
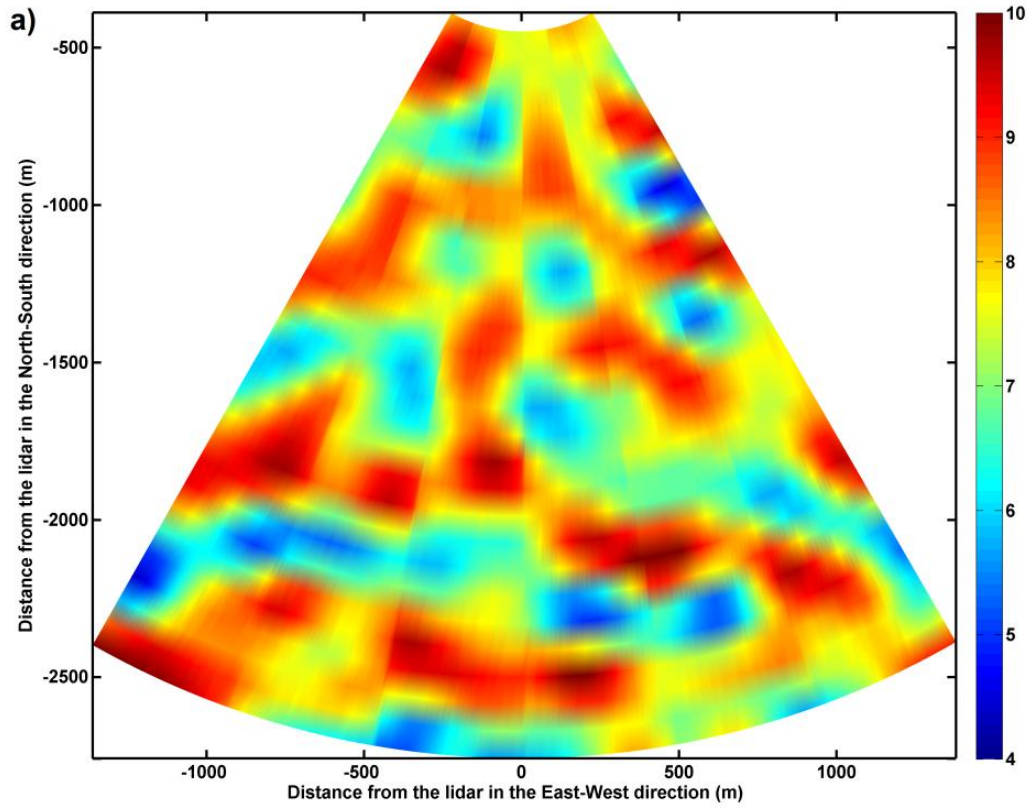

a)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4
b)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 14. Lidar measurement of the flow with scales of motion between 50 m and 150 m and look angle close to perpendicular. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion between 50 m and 150 m with proper truncation number. The colors indicate magnitude of wind speed.


a)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

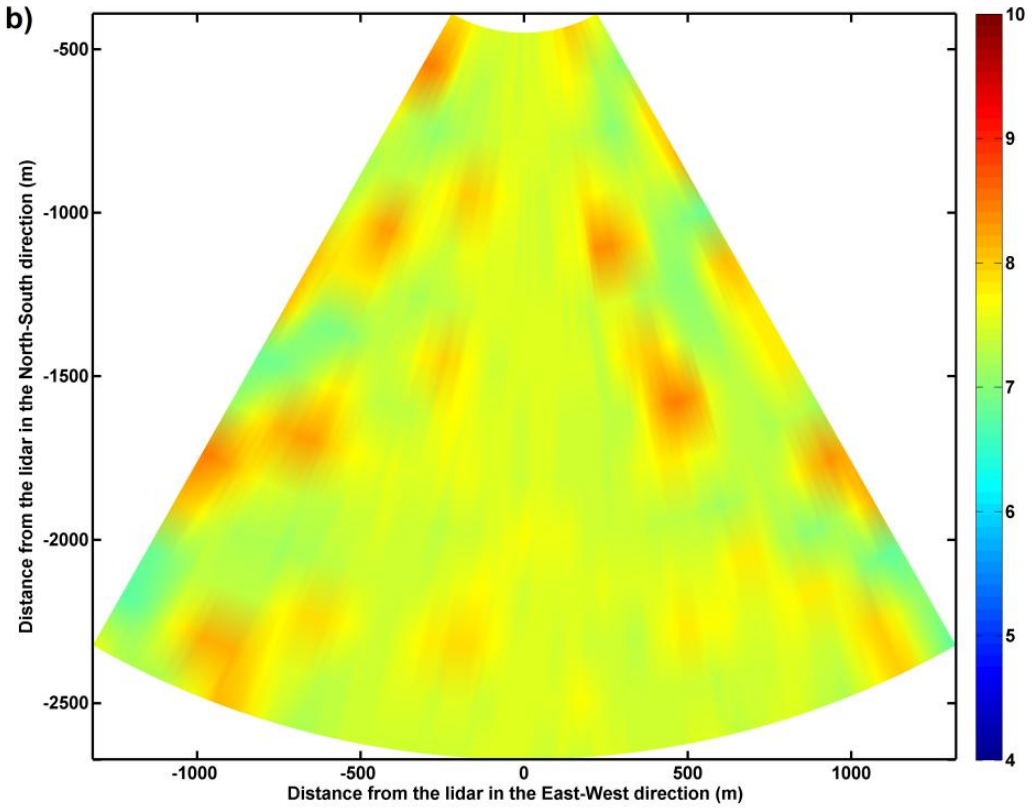

b)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 15. Lidar measurement of the flow with scales of motion between 150 m and 300 m and look angle close to perpendicular. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion between 150 m and 300 m with proper truncation number. The colors indicate magnitude of wind speed.

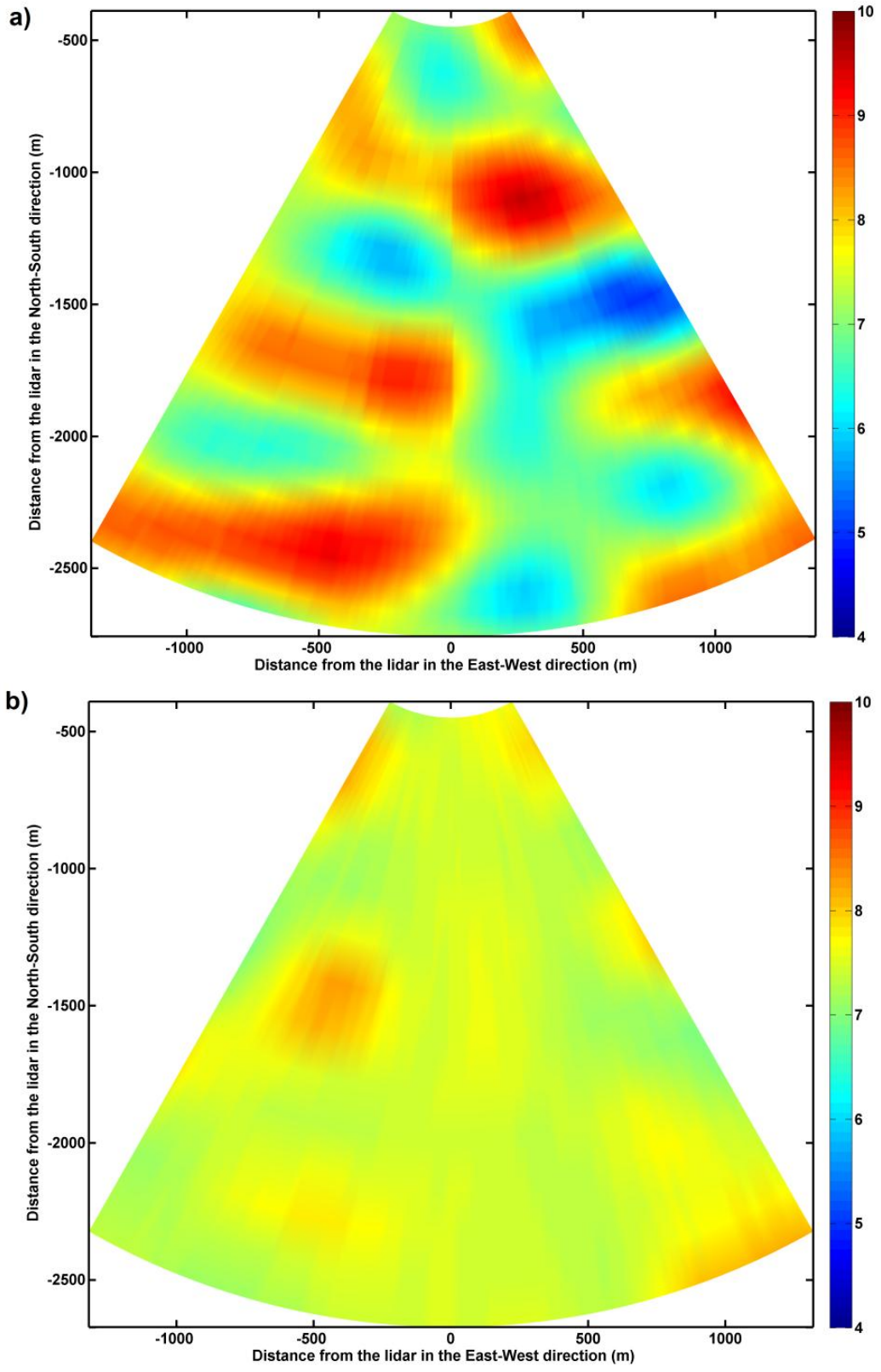

a)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4
b)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-1000
-500
0
500
1000
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 16. Lidar measurement of the flow with scales of motion above 300 m and look angle close to perpendicular. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion above 300 m with proper truncation number. The colors indicate magnitude of wind speed.

Figure 15(a) shows the scales of motion between 150 m and 300 m as observed by the lidar with the look angle perpendicular to the wind speed direction. It is observed that these scales of motion are better represented as was the case with the lidar looking into the wind. The vector retrievals on these scales of motion are shown in Figure 15(b). It is observed that the retrieval accuracy is much lower compared to the case of the lidar looking into the wind. This is due to the fact that the radial velocity does not capture the component of the wind perpendicular to the beam direction. Similar results were observed for scales of motion greater than 300 m as seen in Figure 16.

The vector retrieval accuracy for look angles between 0 degrees and 90 degrees is shown in Figure 17 through Figure 20. It can be observed that as the angle approaches 0 degrees, the flow features are better retrieved versus when the angle is close to 90 degrees, where only the mean flow is preserved.

It should be noted that this analysis was conducted on a flow with an unnaturally high turbulent intensity of 45%. With this in mind, the actual error values are not reported as they would not correspond to actual performance. A LES field with realistic turbulent intensity is being simulated at the National Center of Atmospheric Research (NCAR) and when available will be used to quantify the vector retrieval errors.

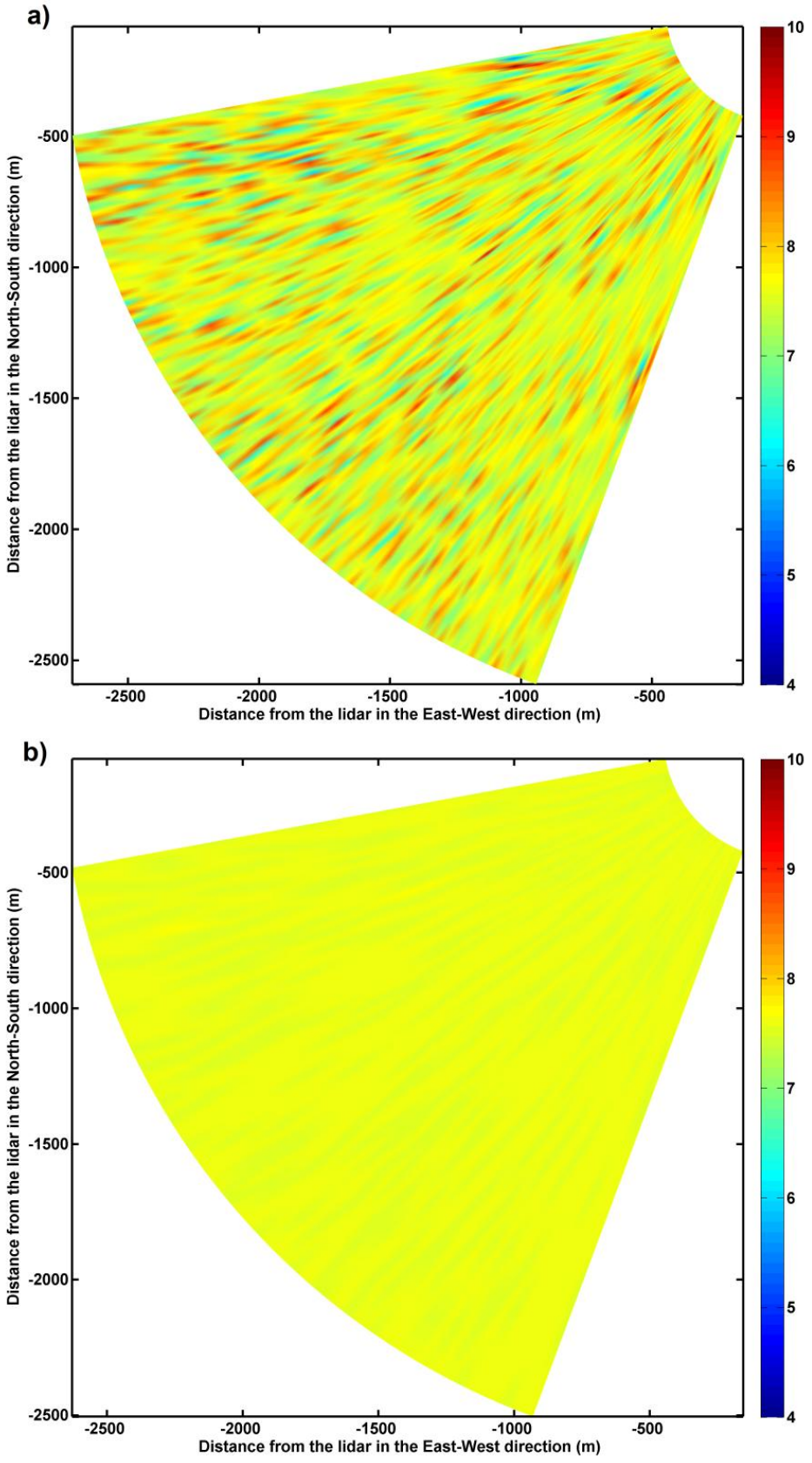

a)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4
b)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 17. Lidar measurement of the flow with scales of motion between 5 m and 50 m and look angle 45 degrees. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion between 5 m and 50 m with proper truncation number. The colors indicate magnitude of wind speed.

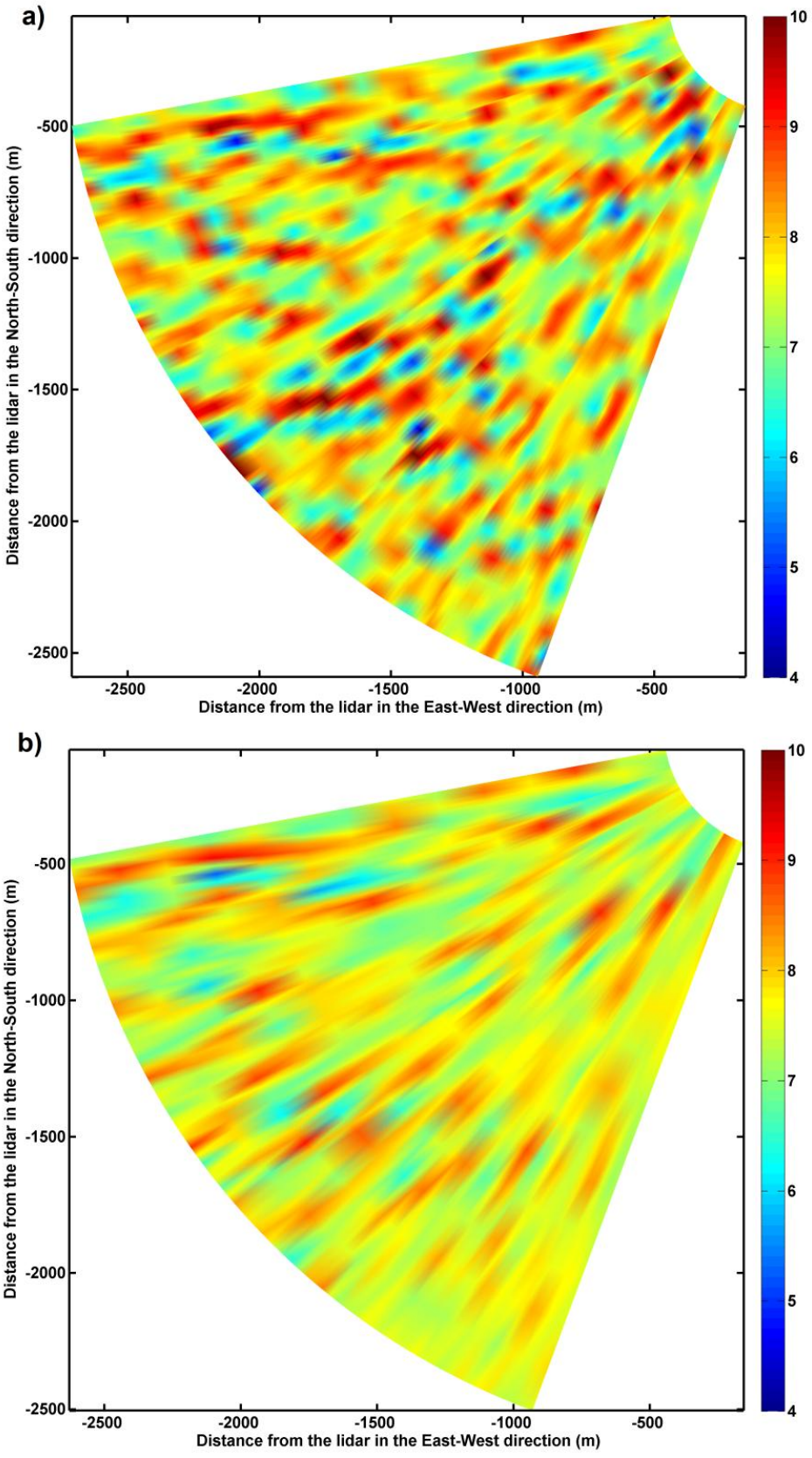
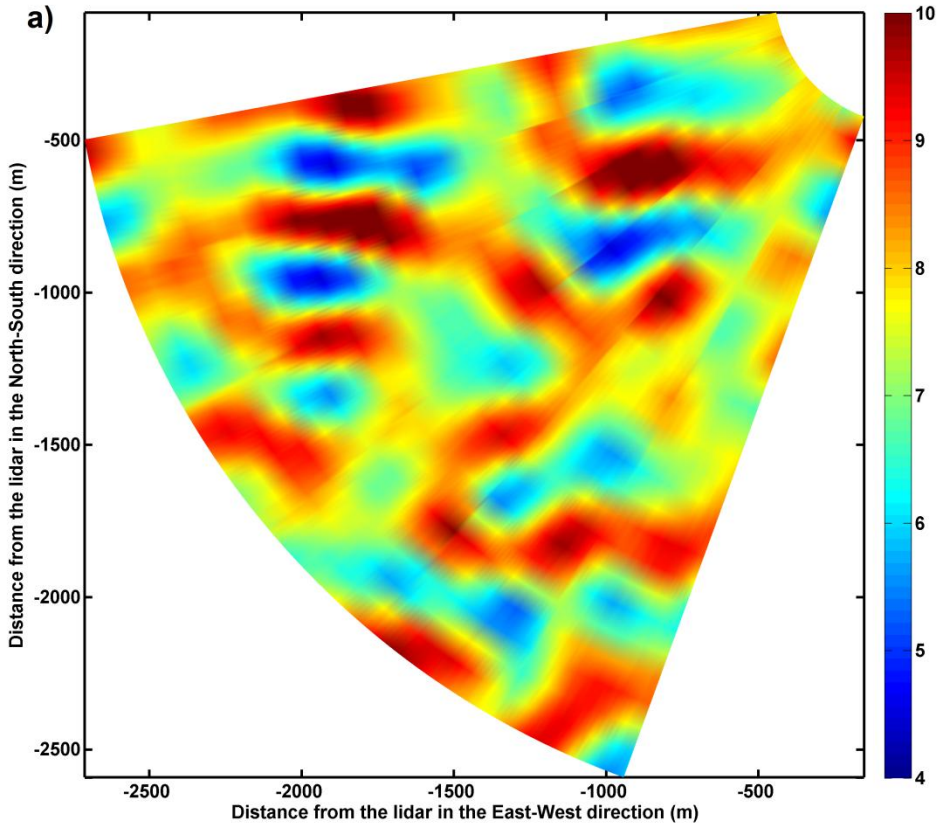

a)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4
b)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 18. Lidar measurement of the flow with scales of motion between 50 m and 150 m and look angle 45 degrees. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion between 50 m and 150 m with proper truncation number. The colors indicate magnitude of wind speed.


a)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

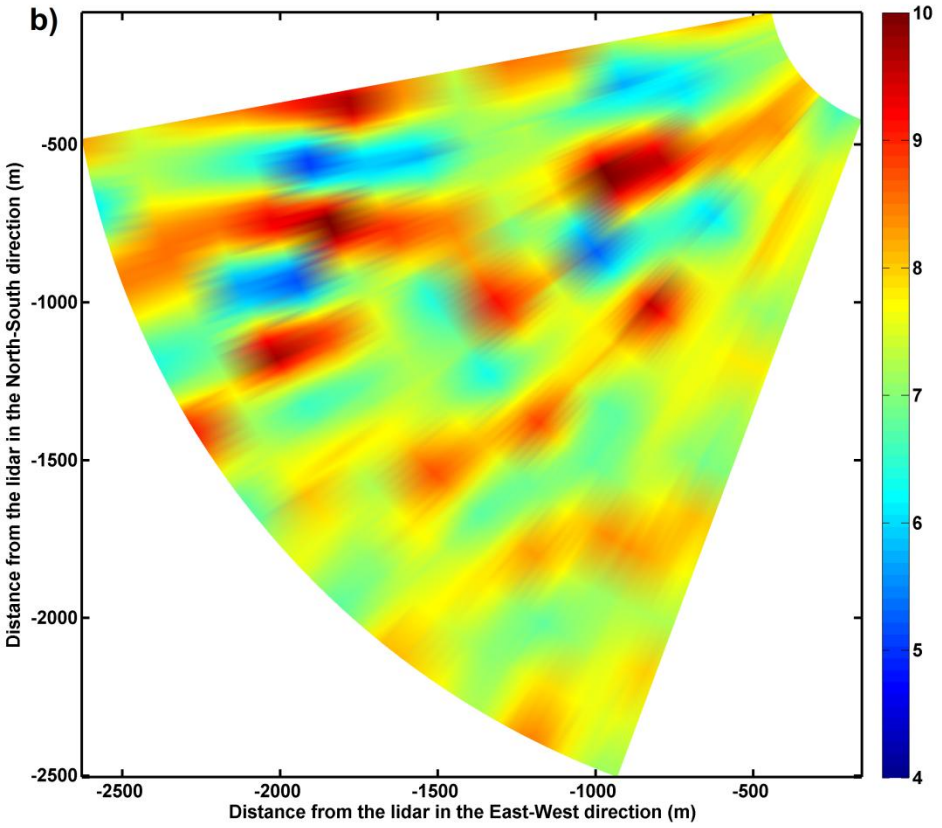

b)
Distance from the lidar in the North-South direction (m)
-500
-1000
-1500
-2000
-2500
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 19. Lidar measurement of the flow with scales of motion between 150 m and 300 m and look angle 45 degrees. (a) The “truth” as seen by the lidar. (b) OI retrieval of scales of motion between 150 m and 300 m with proper truncation number. The colors indicate magnitude of wind speed.

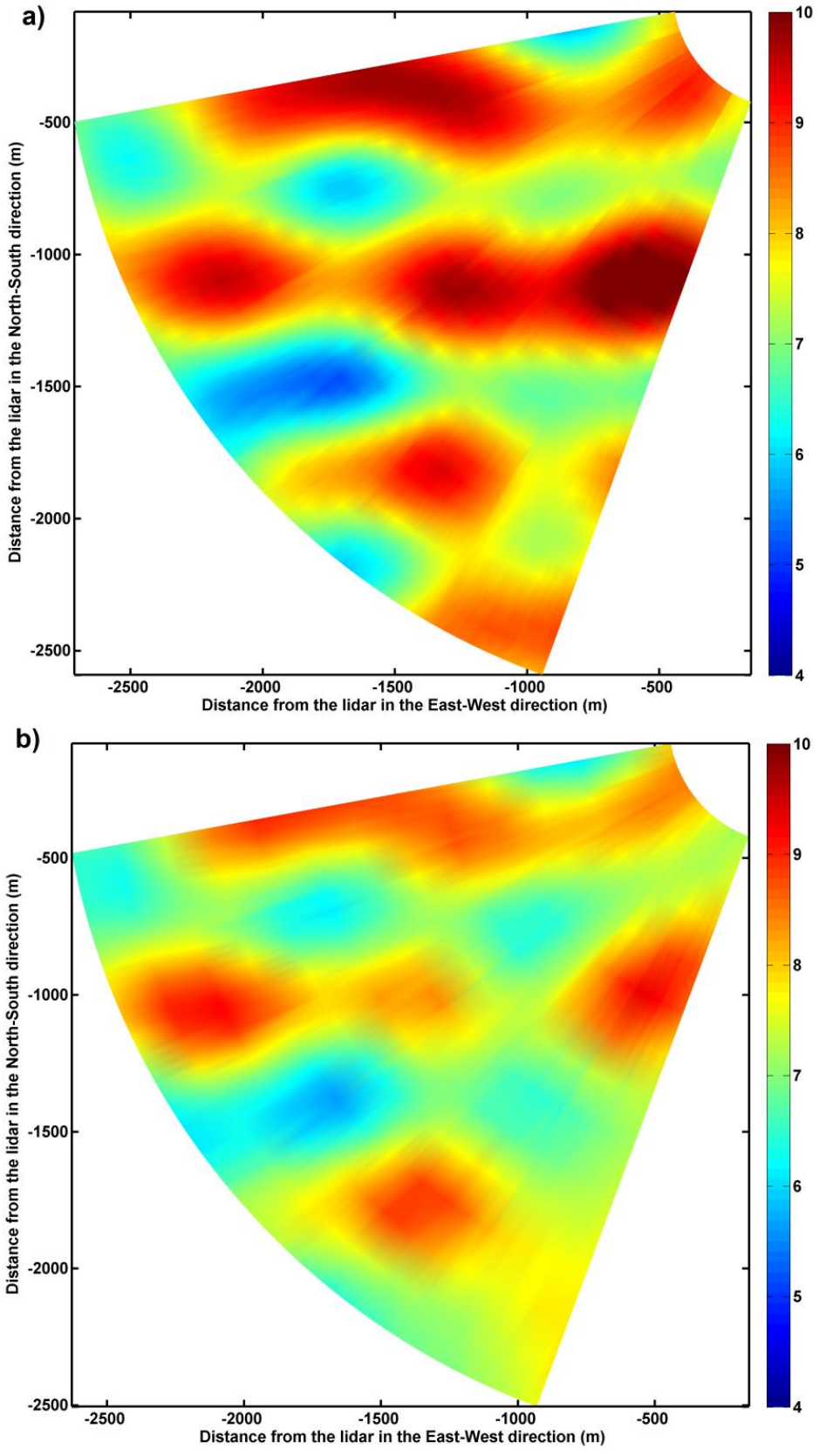

a)
b)
-500
-1000
-1500
-2000
-2500
Distance from the lidar in the North-South direction (m)
-2500
-2000
-1500
-1000
-500
Distance from the lidar in the East-West direction (m)
10
9
8
7
6
5
4

Figure 20. Lidar measurement of the flow with scales of motion above 300 m and look angle 45 degrees. (a) The "truth" as seen by the lidar. (b) OI retrieval of scales of motion above 300 m with proper truncation number. The colors indicate magnitude of wind speed.

4.8 Conclusions

A detailed analysis of the error of representativeness and vector retrieval error was performed. Sensitivity analysis of these errors to lidar look angle and scales of motion present in the flow was carried out. It was found that for scales of motion below 150 m, the error of representativeness of lidar measurement was high. The reason for this is the range-gate size of 80 m which results in under-sampling of the flow field. This effect is akin to the aliasing problem in signal processing. As a result, the lidar simulator measurements showed unnatural streaks along the radial direction formed due to the effect of the range-gate weighting function. These unnatural streaks were removed with proper selection of truncation number in the modified OI technique.

Retrieval accuracy was seen to degrade as the look angle approached 90 degrees to the flow direction. It is observed that for look angle perpendicular to flow direction, only the mean wind speed is retrieved and the local variations along the perpendicular direction are lost. It is observed that a look angle close to 45 degrees provides the optimal retrieval accuracy of both the u and v components.

Future work will entail performing this analysis on flow with a more reasonable turbulent intensity (TI). With a normal TI case, the vector retrieval error and the error of representativeness as a function of scales of motion and look angle can be quantified. In

addition, the analysis will explore the effects of other scanning parameters such as scanning speed, pulse width and elevation angle. Presently, results only for a single range-gate size and zero degree elevation angle are presented. It is expected that non-zero elevation angle will create complex interaction effects between the horizontal and vertical extents of the scales of motion. These effects will be studied once the LES result with normal TI values is available.

5. INVESTIGATION OF A COMPLEX NOCTURNAL FLOW IN OWENS VALLEY, CALIFORNIA

5.1 Introduction:

Flows in mountain valleys display features over a wide range of scales due to motions generated by interaction with the terrain and mountain slopes. An understanding of the fluid mechanics behind such flows would aid in a variety of applications, for example, dispersion modeling of pollutants and particulate matter in valleys, or wind farm characterization in complex terrain. In addition, powerful vortical flows (rotors) formed in mountainous terrain have been linked to numerous aeronautical accidents (Kahn et al. 1997; Doyle and Durran 2002) and therefore are relevant for the airline industry.

Flow over complex terrain, especially mountain valley winds depend on several factors such as horizontal pressure gradients, thermal differences that arise due to topographic amplification factor (TAF – Whiteman 1990), and valley geometry. Over the years extensive research has led to improved understanding of meteorological processes in mountain valleys. There have been a number of studies, both observational (for example: Whiteman et al. 2009; Princevac et al. 2008) and numerical (for example: Jiang and Doyle 2008; Grubišić and Billings 2007) which have investigated thermally driven flows, mountain waves, temperature variations and diurnal effects within mountain valleys.

Whiteman and Doran (1993) with the help of climatology and numerical investigations determined that flows within a valley are affected by locally developed

along-valley pressure gradients, downward momentum transport and pressure-driven channeling of the geostrophic wind into the valley. Flows in mountain valleys are seen to have strong diurnal characteristics. Generally flow during the daytime is up-valley and up-slope, while nocturnal winds are down-valley and down-slope (Whiteman 1990). Often flows in mountain valleys deviate from this ideally defined behavior. For example, downward momentum transport and pressure-driven channeling can affect the normal diurnal cycle of the valley winds. The degree of disruption in the valley wind system depends on several conditions such as the stability of the boundary layer within the valley, strength of the geostrophic wind, and channeling of the winds by the valley walls (Whiteman 2000). Schmidli et al. (2009) describe two non-classical flow events due to thermally driven night-time flows: 1) the evolution of a three layer flow with an up-valley winds sandwiched between a thermally driven down-valley flow and a synoptically driven north-westerly flow and 2) a strong and deep down-valley jet with speeds in excess of 15 m s-1. Using observations and model results from the Advanced Regional Prediction System (ARPS), a three layer flow structure observed within the Owens Valley is described by Schmidli et al. (2009). Princevac et al. (2008) describe a theoretical and observational study of nocturnal katabatic flows in wide valleys with weak synoptic winds. It is reported that katabatic winds exhibit along-slope periodic oscillations resulting in sustained turbulence in nocturnal flows. These night time oscillations were also observed in previous studies (see Porch et al. 1991) where the oscillations of drainage flows are reported resulting in cold air intrusions into the valley. Similar deviations from expected behavior do occur frequently (Whiteman 2000) and further observation and analysis is required to understand them completely.

Lidar remote sensing offers a valuable approach for conducting these types of studies due to its ability to obtain observations with high enough spatial and temporal resolution to capture transient fluid dynamical features (see, for example; Banta et al. 1999). In addition, the ability of Doppler lidar to perform rapid, configurable scans allows study of three-dimensional flow structures. For example, there have been several studies that made use of Doppler lidar to study the flow on vertical cross-barrier planes during the Terrain-induced Rotors EXperiment (T-REX) (see Hill et al. 2010 and Weissmann et al. 2009). Hill et al. (2010) made use of intersecting vertical lidar scans to track and characterize mountain-induced rotors and other flow structures in clear air. Weissmann et al. (2009) calculated the vorticity in the vertical two-dimensional plane using continuity to derive the velocity component perpendicular to the line of sight. Banta et al. (2002) and Banta et al. (2008) describe complex flow phenomena such as nocturnal jets and periodic shear instabilities captured using lidar remote sensing methods. In this paper, a vector retrieval method based on optimal interpolation, recently adapted from the radar processing community for Doppler lidar (see Xu et al. 2006; Kongara et al. 2011), is used to educe the velocity vectors on horizontal lidar scans and study a complex flow event observed during the early morning of March 27 2006 in Owens Valley, California. In addition, results from the Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS®) are used to understand the mesoscale flow structure.

5.2 Experimental Setup

The Terrain Induced Rotor Experiment (T-REX) was a study carried out in the Owens Valley of the Sierra Nevada mountain range in the state of California (USA) to

study the formation of rotors and study thermally-forced flows in complex terrain. A detailed description of the T-REX campaign is given by Grubišić et al. (2008). The main objective of this study was to improve the understanding and predictability of mountain-waves, rotors, and boundary layer flow in complex terrain. This campaign was carried out for a period of two months from March 2006 to April 2006. The terrain surrounding the Owens Valley is complex and known to produce rotors and recirculations in the atmosphere. The Owens Valley is approximately 100 km long and is oriented approximately North-Northwest and South-Southeast. It is surrounded by the Sierra Nevada Mountains on the west and the Inyo and the White mountains on the east. The majority of the instrumentation was set up near the town of Independence, California (36°48'9.76"N, 118°12'0.34"W). The valley floor is approximately ten kilometers wide near Independence. This area is known for atmospheric rotors, potentially dangerous for aircraft, and strong winds which can lead to high aerosol events.

A network of in situ, ground-based and airborne instruments was deployed by various organizations (Grubišić et al. 2008) to make measurements of the flow structure in the valley. Among the ground-based instruments were two Doppler lidars, one from Arizona State University (ASU) and a second one from the Deutsches Zentrum für Luft- und Raumfarht (DLR, German Aerospace Centre) both WindTracer type, built by Lockheed Martin Coherent Technologies, Inc. Both lidars had a similar configuration with a 2-μm laser beam, a pulse repetition frequency of 500 Hz and a range-gate size set to 72 m. The two lidars were placed approximately 2.9 km apart and performed scans which frequently overlapped, thus allowing the possibility of dual Doppler analysis. The DLR lidar was located at 36°47'43.33"N, 118°12'28.74"W at an elevation of 1242 m

above MSL (mean sea level) and the ASU lidar was located at 36°47'51.74"N, 118°10'32.06"W at an elevation of 1179 m above msl. The positions of the lidars and other instrumentation used in this analysis are shown in Figure 21. The difference in elevation of the two lidars is 63 m. The other instruments whose data were used in this paper are: 1) the University of Leeds Radiosonde (released at 118o 12' W, 36o 48.60 'N), 2) the National Centre for Atmospheric Research (NCAR) Integrated Surface Flux Facility (ISFF) at the Central tower (See Fig. 1 for location), and 3) NCAR Mobile Integrated Sounding System (MISS) Radio Acoustic Sounding System (RASS) located at 36o 47.2' N, 118o 12' W during this time period.

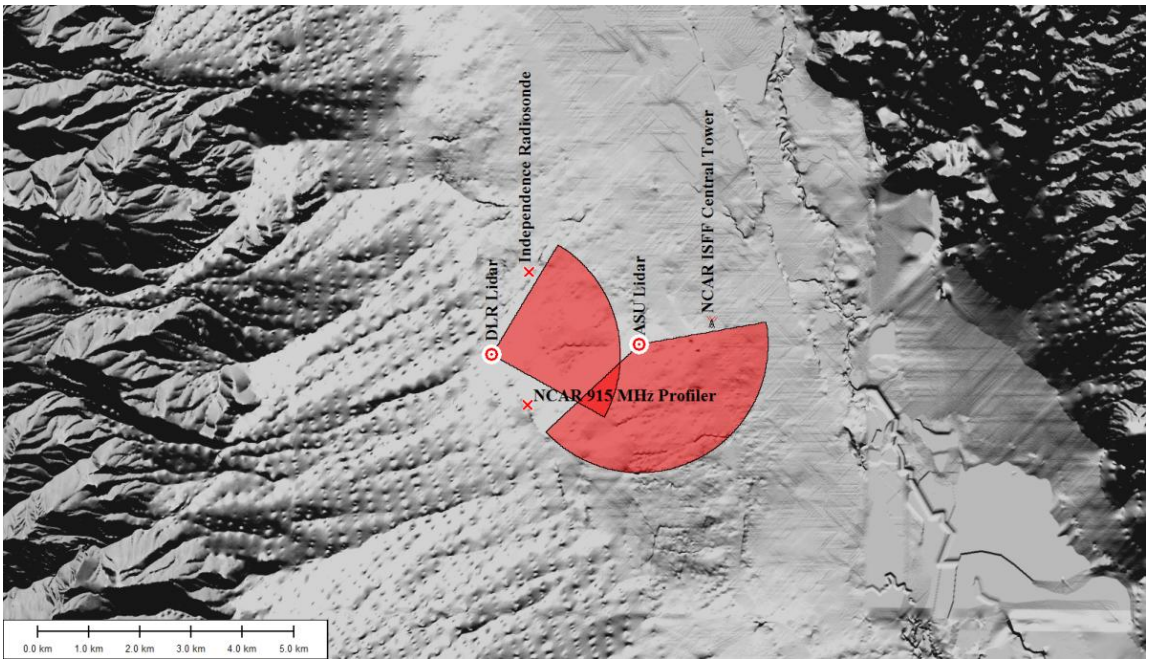


Figure 21. Locations of the instruments used in the analysis. The red areas are the scan regions of the two lidars.

In this study, data collected by the lidars on March 27 2006 are investigated. On this day both the ASU and the DLR lidar executed a stack of PPI (Plan Position

Indicator) scans at elevation angles ranging from 0^o to 26^o with horizontal steps of 2^o. The DLR lidar scanned from an azimuth angle of 31^o to 119^o, measured clockwise from north, while the ASU lidar scanned from 127^o to 219^o, measured clockwise from north. Here, particular focus will be given to observations made during the time period of 1030 to 1230 UTC (0230 to 0430 hr local time) early morning of March 27, 2006. This time period was selected because of complex, converging flow patterns observed in the Doppler lidar data. It is observed that there was a change in the direction of flow with height with strengthening and weakening of the down-slope winds and evidence of a cross-valley flow impinging from the west.

5.3 Discussion

The flow in the Owens Valley during the early morning of March 27 2006, displays several interesting features. In this paper, the flow characteristics observed between 0900 UTC to 1230 UTC (0100 hrs to 0430 hrs local time) are studied in detail. From lidar observations, it is seen that the flow even after sunset on Mar 26th remains up-valley for the complete depth of the valley. At about 1000 UTC on March 27 (0200 hr local time), a westerly cross-valley flow begins to form in the lower levels of the valley. This westerly cross-valley flow is subsumed in to a down-valley flow that is seen to begin after 1030 UTC (0230 hrs local time). The depth of this down-valley flow is initially quite shallow (eventually reaching a maximum height of about 200 m) and weak with speeds approximately equal to 2-4 m s^{-1}. Interesting flow characteristics are observed starting from 1030 UTC to 1230 UTC (0230 to 0430 hr local time) which will be described below.

In order to interpret the results from the lidar observations, it is necessary to understand the boundary layer stability during this time. The virtual temperature profiles measured by the NCAR MISS RASS and the NCAR ISFF Central Tower within the valley were plotted and are given in Figure 22. Comparing these profiles to the dry adiabatic lapse rate, it can be seen that the boundary layer was: i) Strongly stable (warm air above colder air) for the lowest levels measured by the tower (except at 0930 UTC), and ii) Weakly stable above 200 m.

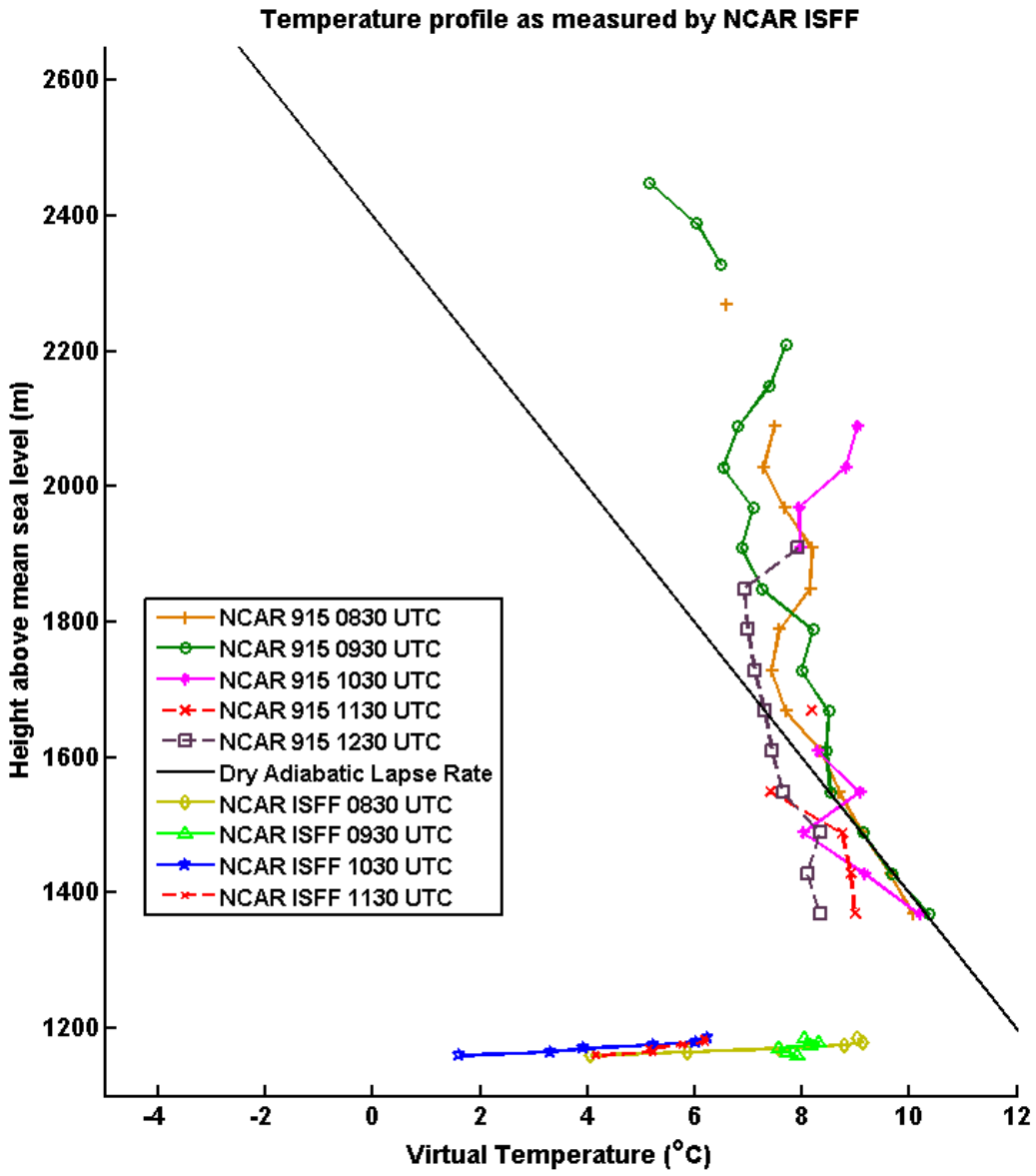


Figure 22. Virtual temperature profiles from 0830 UTC to 1130 UTC on March 27 2006, measured by the NCAR DBS 915 MHz profiler and NCAR ISFF tower. The solid black line shows the dry adiabatic lapse rate.

A high resolution COAMPS run with spatial resolution of 333 m and the radiation model updated every hour was made for March 27 with output every 15 mins during the study period of 0900 UTC to 1230 UTC (0100 hrs to 0430 hrs local time). The high

resolution COAMPS results show that the air above the ridge line (>3000 m above ground level) is south-westerly (see Figure 23 for COAMPS results on a surface 5000 m above ground level) with a speed approximately equal to 10-12 m s^{-1}. Measurements from University of Leeds Radiosonde are used to validate the results from the COAMPS run. It is seen that the COAMPS results agree reasonably with the radiosonde measurements. The first 24 levels of the COAMPS horizontal wind speed and directions at 1500 UTC are compared with the radiosonde measurements at 1537 UTC. The r.m.s. difference between the radiosonde wind measurements and the COAMPS results is 3.1 m s^{-1} and the r.m.s. difference between the wind directions is 13.31 degrees.

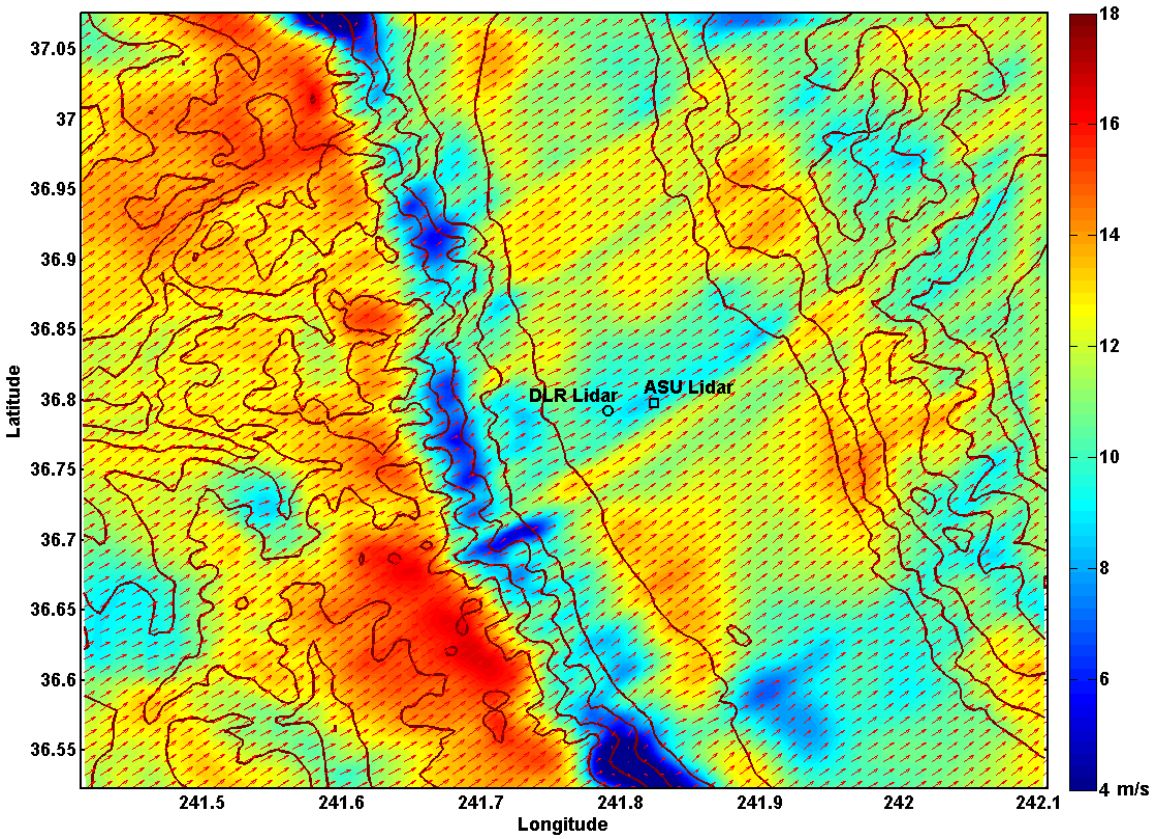


Figure 23. COAMPS winds on a 5000 m height surface at 1030 UTC (0230 hr local time) on March 27 2006. Colors designate the magnitude of the velocity vectors.

It is observed that there is a sudden change in flow dynamics during the time period under consideration. The effect of these flow dynamics on the turbulence levels can be estimated using the gradient Richardson number (Eq. 50), which is the ratio of buoyant destruction of turbulence to the shear production. The 30-minute averaged gradient Ri number profiles generated using the NCAR ISFF data are shown in Figure 24. The gradient Ri number values are between -1 to 1.2 (except at 1030 UTC and 1200 UTC) indicating that the flow is dynamically unstable (close to critical Ri number of 0.25). That is, the turbulence created by the velocity shear is not completely overcome by buoyancy destruction. A sharp jump in the Ri number is observed at 1030 UTC and 1200 UTC. This indicates a sudden increase in stability, which could be brought about either by an increase in the stratification or a reduction in the velocity shear. The reasons for this will be discussed in section 5.4.

$$Ri = \frac{\frac{g}{T_v}\frac{\partial \theta_v}{\partial z}}{\left(\frac{\partial U}{\partial z}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2} \tag{50}$$

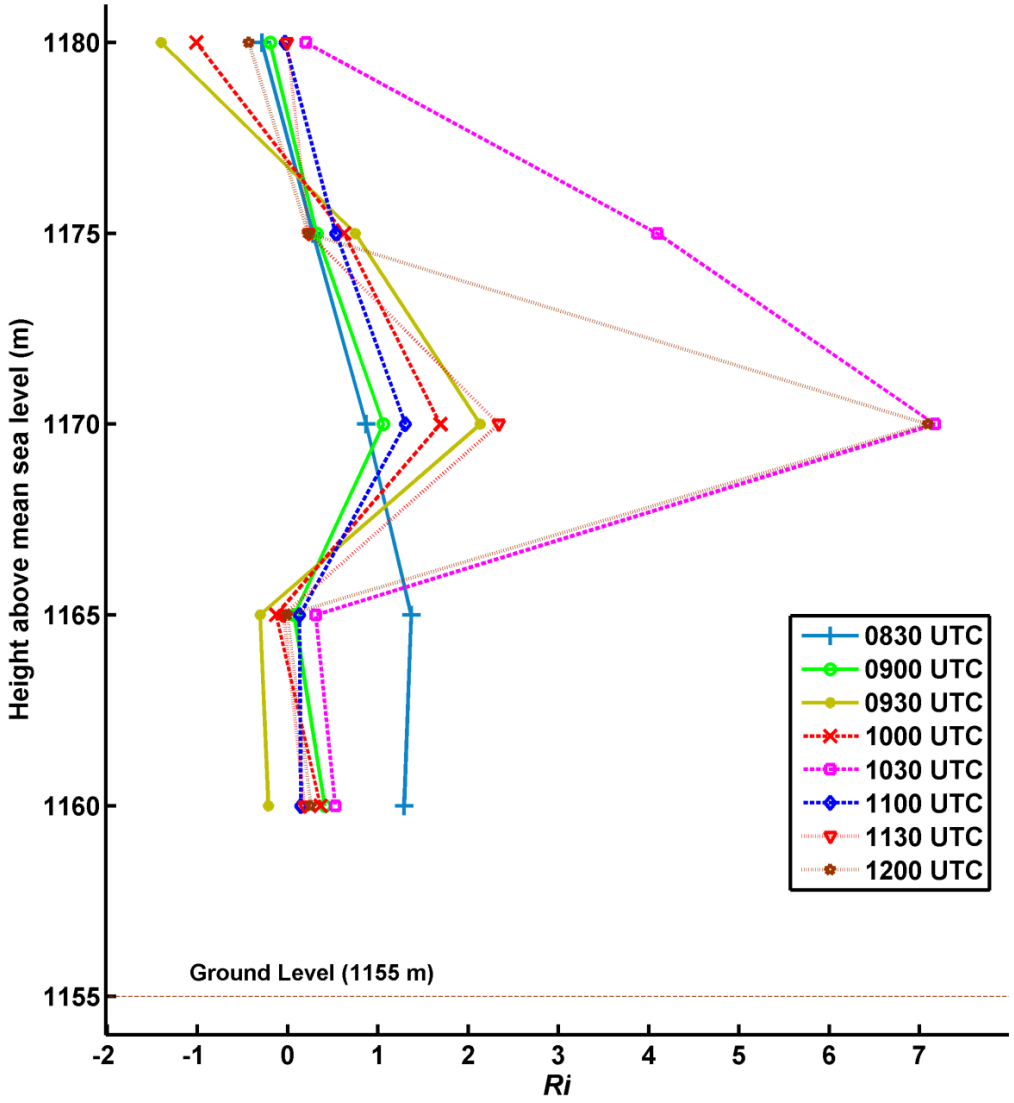


Figure 24. Gradient Ri number profiles in the Owens Valley estimated using data from NCAR ISFF Central tower.

5.4 Lidar observations

The DLR lidar was located near the west valley wall and scanned in the cross-valley direction. Figure 25 shows the retrievals from the DLR lidar scans of zero degree elevation. Figure 25a shows that at 1030 UTC (0230 hr local time), there is a westerly cross-valley current from the west mixing with the down-valley current prevalent at this

level within the valley. It was observed from the DLR lidar measurements that this westerly cross-valley flow began approximately at 1000 UTC. With time, this westerly cross-valley wind gets weaker and is subsumed into the generally down-valley flow, as seen in Figure 25b to Figure 25d. Before this time, the flow was generally up-valley (not shown) throughout the depth of the valley. Near 1030 UTC, the down-valley flow mixes with pulses of cross-barrier momentum flowing down the valley walls.

The ASU lidar located within the valley is well suited to capture the central valley flow. Observations from the ASU lidar also show presence of a westerly cross-valley flow within the valley at approximately the same time as the DLR lidar. The vectors on the 0 degree elevation scan as measured by the ASU lidar are shown in Figure 26. One can see the westerly cross valley flow clearly until 1030 UTC and then it gradually dies out and is replaced by a generally down-valley flow on the lowest levels. These observations also confirm that the westerly cross-valley flow is observed only for 30 minutes (from 1010 UTC to 1040 UTC). Another observation from the ASU lidar is that the cross valley flow is double layered – that is there exists a flow moving towards the east and another toward the west, presumably associated with flows down the valley walls into this particular vertical layer measured by the ASU lidar (60 meters below the DRL). This can be seen from the plot of the velocity vectors measured by the ASU lidar shown in Figure 27. It is seen from the 0^o elevation scan (Figure 27a) that the flow is from the west to the east. On the 1^o elevation scan, one can see that the flow closer to the lidar (at lower elevation) is directed west-to-east, while at farther range-gates (higher elevation), the direction is from east-to-west.

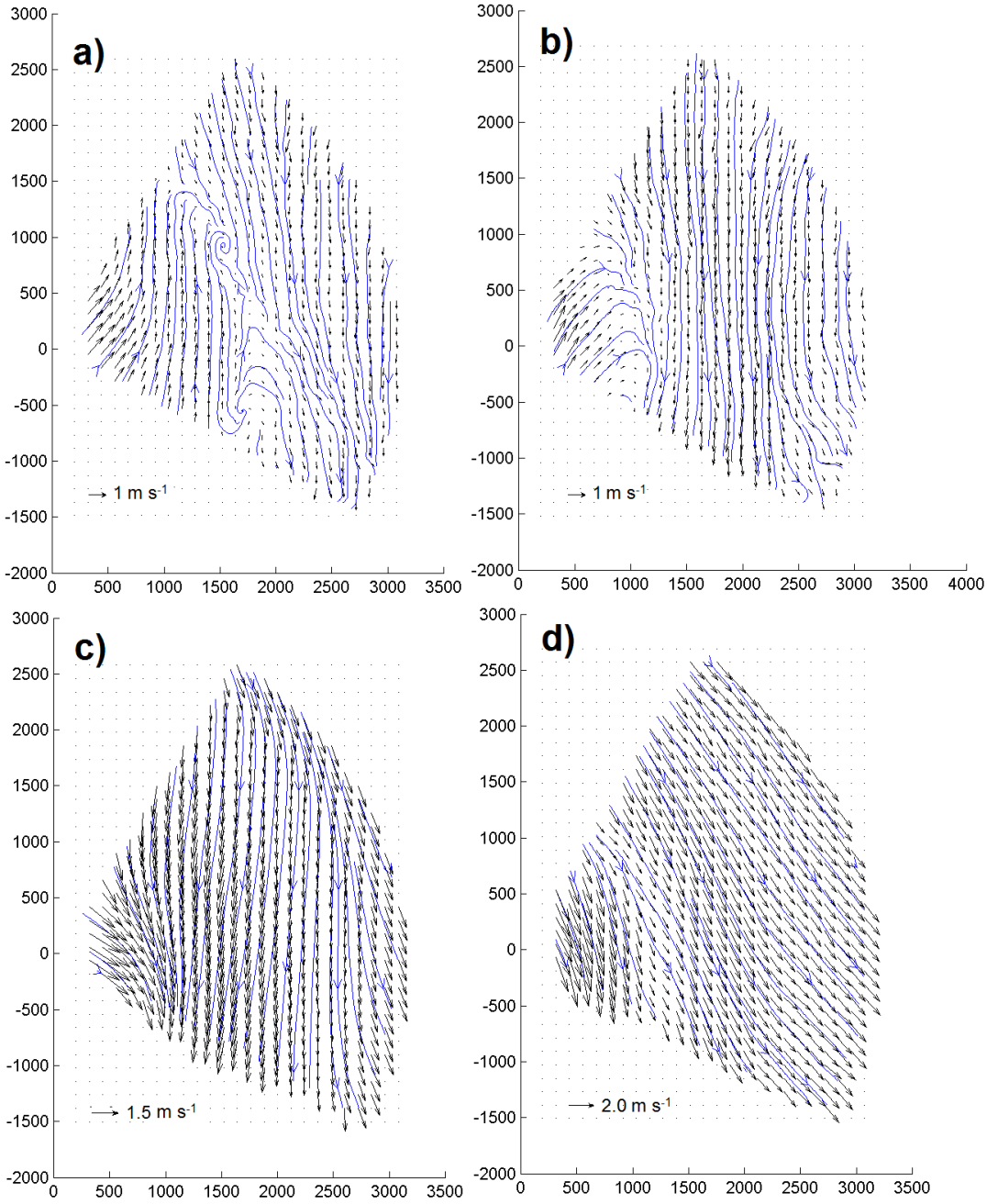


Figure 25. Velocity retrievals from the zero degree elevation DLR lidar scan. a) 1030 UTC (0230 hr local time) b) 1050 UTC (0250 hr local time) c) 1100 UTC (0300 hr local time) d) 1105 UTC (0305 hr local time).

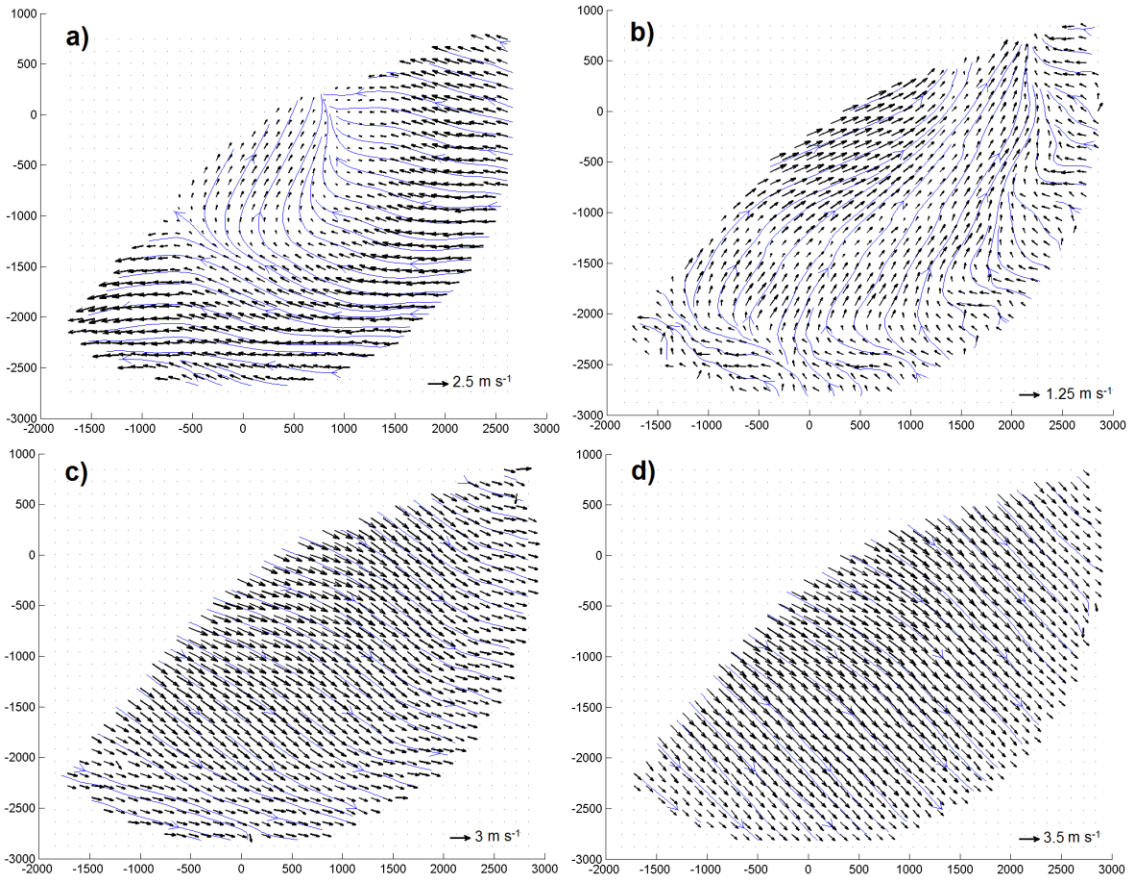


Figure 26. Velocity retrievals from the zero degree elevation ASU lidar scan. a) 0940 UTC (0140 hr local time) b) 1030 UTC (0230 hr local time) c) 1050 UTC (0250 hr local time) d) 1120 UTC (0320 hr local time).

Similarly, Figure 27c and Figure 27d show the flow at higher elevations is from east-to-west. Such abrupt changes in wind speed and direction at low heights (50 – 200 m as observed here) are accompanied by high wind shear, and therefore are potentially relevant for understanding valley pollution scenarios or for applications such as wind farm design in complex terrain. Figure 28 shows a comparison of 15 minute averaged vertical profile of wind speed and wind direction obtained from the ASU and DLR lidar at 1030 UTC (0230 hrs local time). The vertical profiles from COAMPS results are also

shown. Due to data availability considerations, the distance between the locations of the lidar profiles is approximately 1.1 km. The COAMPS profile is plotted at a point approximately equidistant from the ASU and DLR lidar profiles. The vertical profiles for the lidars are formed by plotting the wind speeds retrieved from the OI method for the low elevation angle scans, and the VAD for the high elevation angle scans with height. Due to the difference in the nature of the retrieval methods, the spatial averaging is different for the lower and higher levels. On the lower levels, there is no averaging in the azimuthal direction while the averaging is about 1.0 km in the radial direction. Vectors from the first 10-15 range-gates of a given beam are used for binning depending on the elevation angle of the scan. While on the higher level, since VAD is used for vector retrieval, there is spatial averaging in the radial as well as azimuth direction. This spatial averaging on the upper (VAD) levels is of the order 3 km in the azimuthal direction and 1.0 km in the radial direction. It is observed that the profiles of wind speed and wind direction agree reasonably given the spatial distance between them and the spatial homogeneity of the flow. Evidence of the low level westerly cross-valley flow is seen from the ASU lidar wind direction retrieval. The COAMPS results are discussed in Section 5.5.

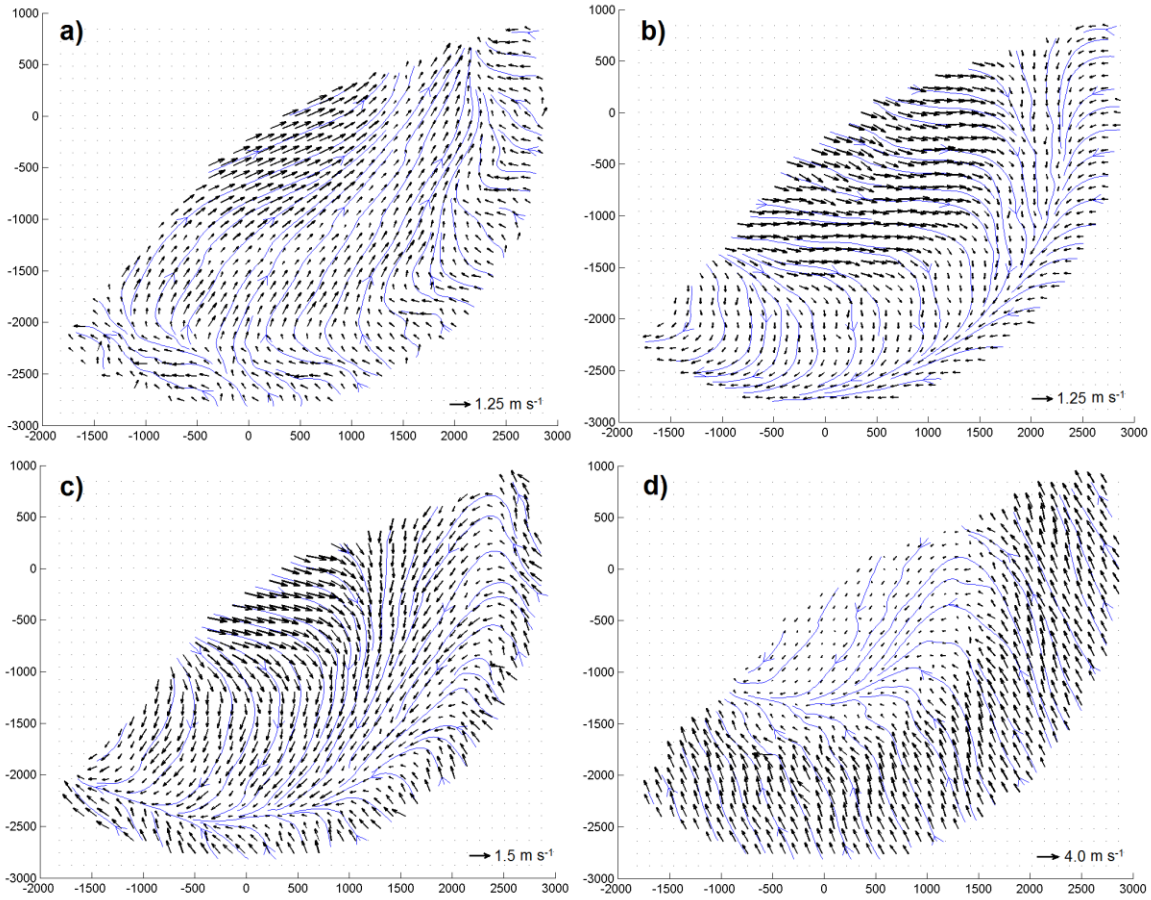


Figure 27. Velocity retrievals from the ASU lidar scan at different elevation angles: a) 0^o elevation angle (10:30:23 UTC) b) 1^o elevation angle (10:31:01 UTC) c) 2^o elevation angle (10:31:39 UTC) d) 4^o elevation angle (10:32:17 UTC).

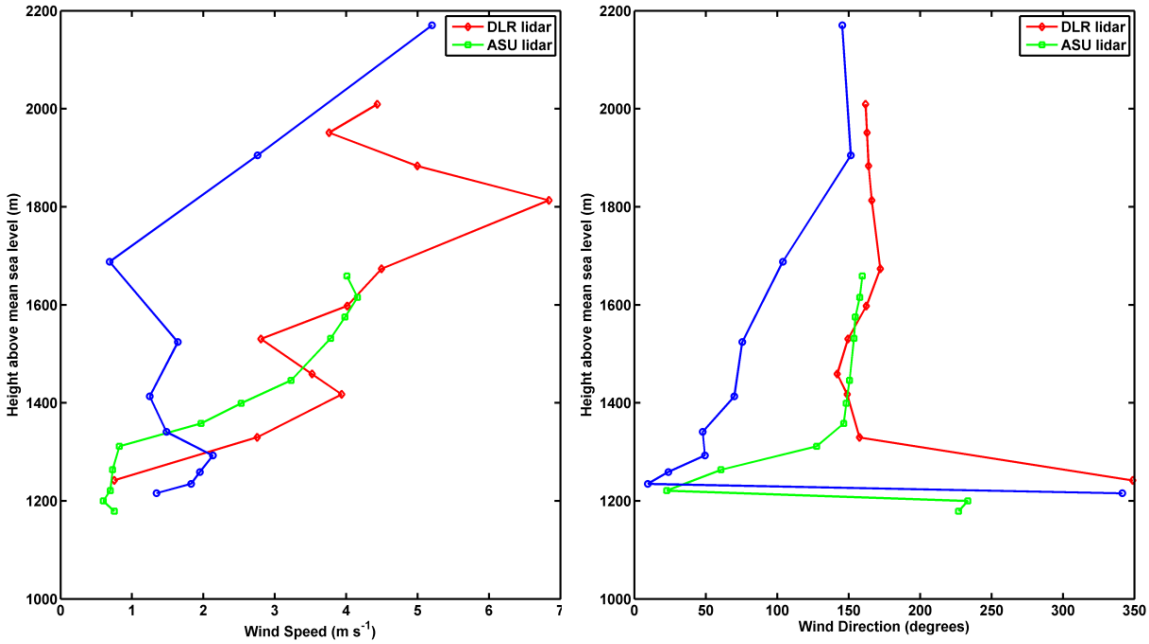


Figure 28. Comparison of the vertical profiles of (a) wind speed and (b) wind direction obtained from the ASU and DLR lidars and COAMPS at 1030 UTC on March 27 2006.

Figure 29 shows 30-minute averaged vertical flow structure as observed from the ASU lidar. This plot made in a similar manner as the previous plot. It can be seen that the lower down-valley flow increases in strength and height with time. This behavior is expected because as the night progresses, the drainage flow down the middle of the valley is fed by sporadic pulses of downslope flows from the valley sides. Consequently, the height of the stably stratified boundary layer increases, strengthening the down-valley flow.

A sharp drop in temperatures near the valley floor is measured by the NCAR ISFF tower during this time (see Figure 30). This steep temperature gradient could be caused by a pulse of cold air intruding onto the lowest levels of the valley flow. These steep temperature drops are seen to occur periodically until sunrise at 1430 UTC and are

accompanied by the low level westerly cross-valley currents. It is observed that there are periods of stable stratification followed by periods of vertical mixing occurring repeatedly until sunrise.

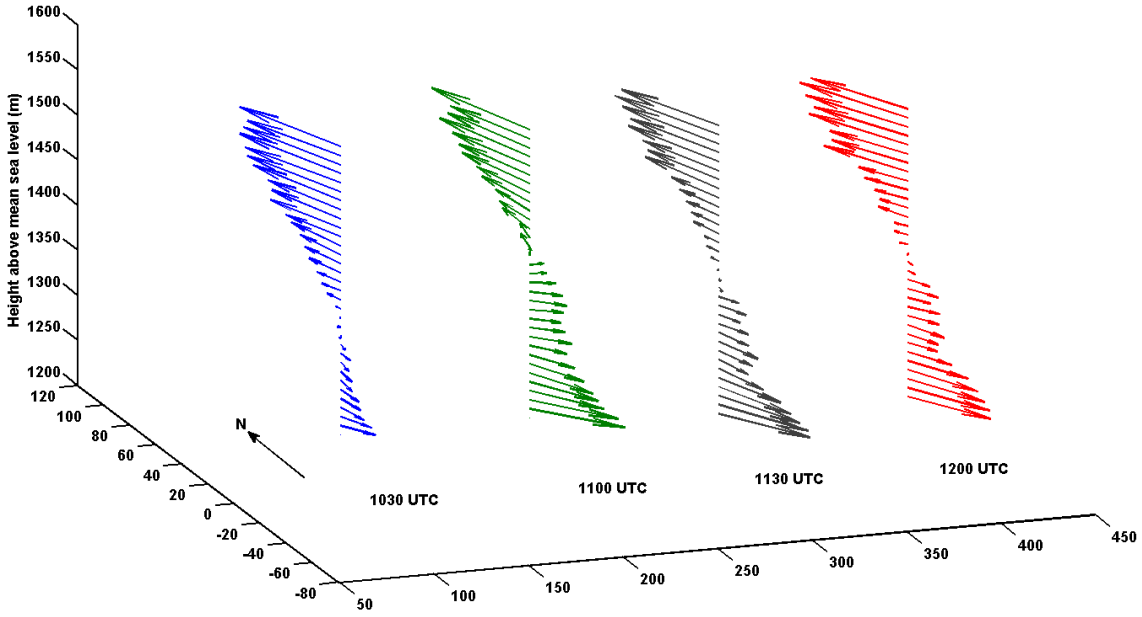


Figure 29. Vertical profiles of 30-minute averaged horizontal velocity vectors from ASU lidar showing the low level down-valley flow and the upper level up-valley flow in the Owens Valley between 1030 UTC to 1230 UTC on March 27 2006.

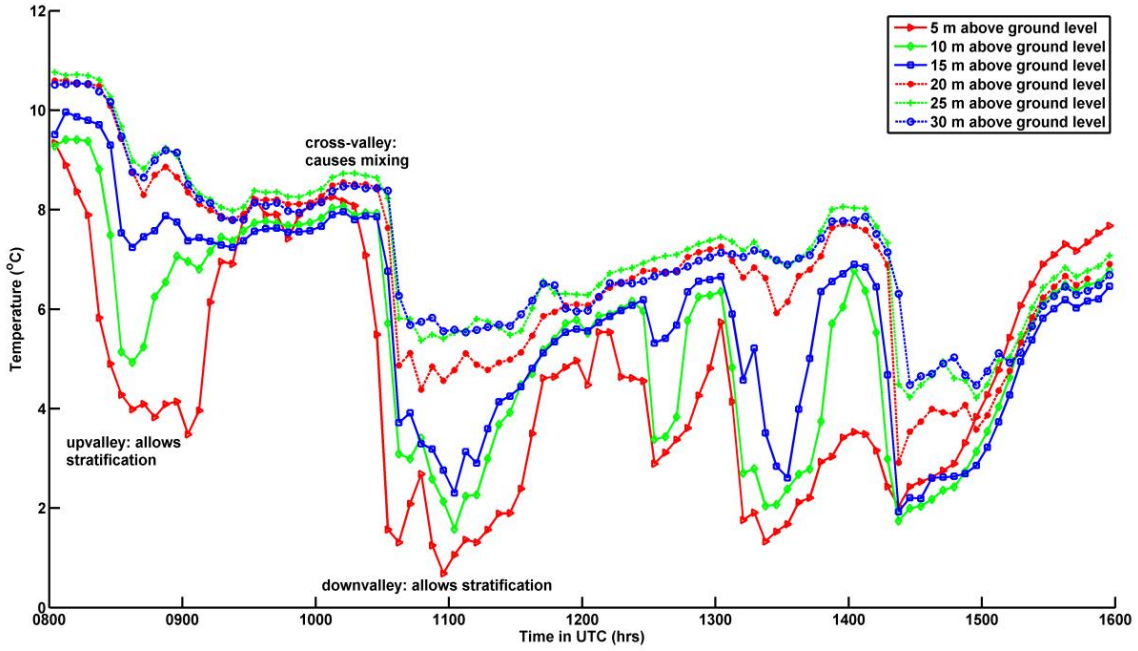


Figure 30. Time series of the virtual temperature measured by the NCAR ISFF tower.

The periodic cross-valley flow events resemble wind break-in events (Whiteman 2000). Usually break-in events occur during daytime, when the convective boundary layer increases in height and re-couples with the winds above the ridgeline. But break-in events can also occur if the boundary layer within the valley is weakly stable as in our case and winds above the ridgeline are strong and flowing in the cross barrier direction. These break-in events can cause sudden changes to the local wind flow and disrupt the diurnal wind system. As can be seen from Figure 30 and Figure 31, the westerly cross-valley flow events cause vertical mixing and destroy the wind shear present. In addition to the tower data, time series from the ASU lidar from points closest to the NCAR ISFF tower location are also plotted (see black curve). The spatial averaging on this data is 100 m and the location where data is plotted is approximately 700 m from the NCAR

ISFF tower location. It is observed that the flow between 0930 UTC and 1000 UTC is cross-valley.

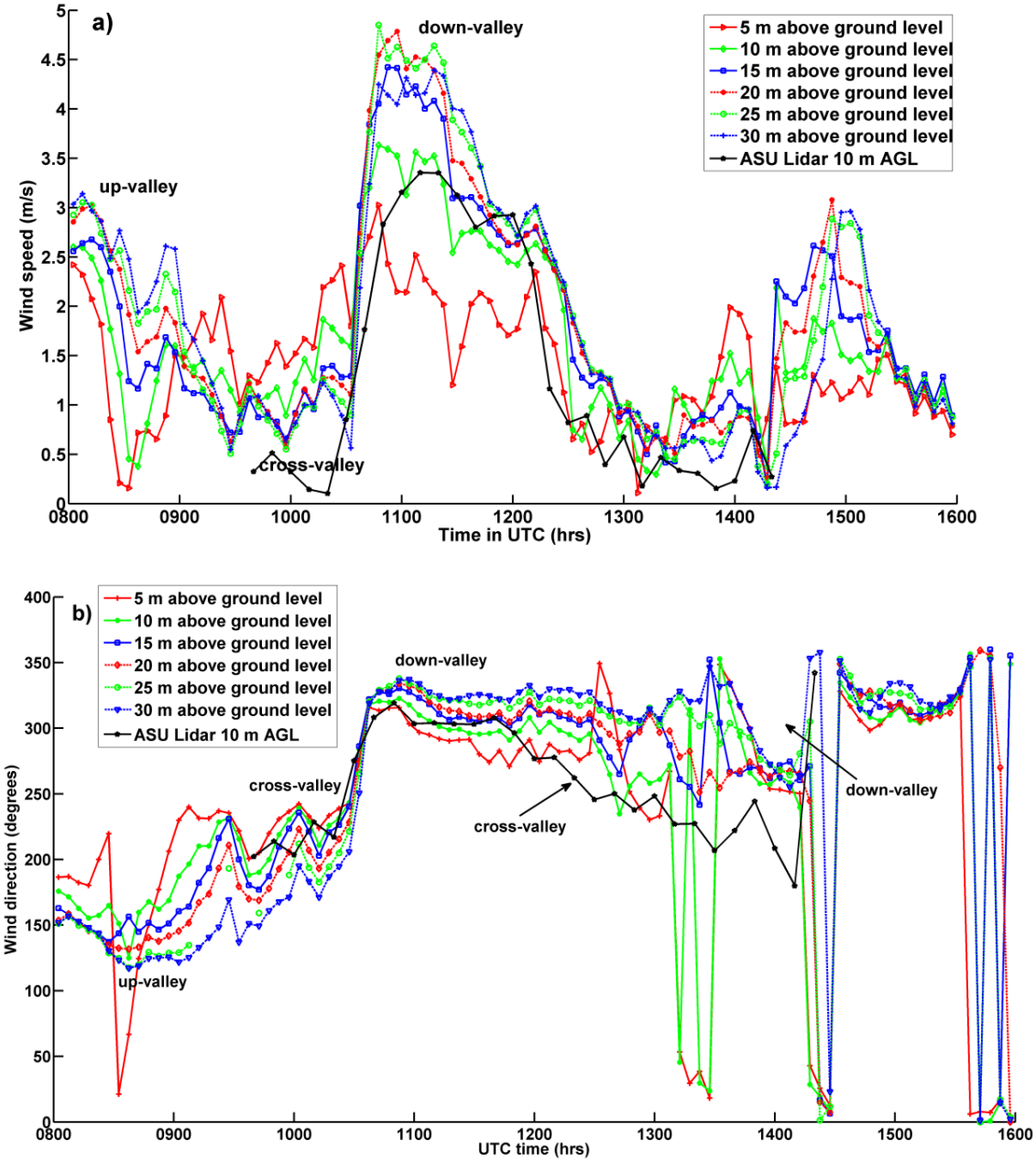


Figure 31. Time series of the (a) wind speed and (b) wind direction measured by the NCAR ISFF and the ASU lidar closest to the ISFF tower location.

During this time, the wind speeds are low and the temperature profiles show vertically mixed conditions in the lowest 30 m of the valley. At around 1030 UTC, the flow is still cross-valley from the west, but thermally stratified conditions exist. This thermally stratified condition exists till about 1200 UTC. Between 1200 UTC and 1400 UTC, the lidar measures the flow to be cross-valley coming from the west. The tower measurements at the same height above ground level agree with this until about 1300 UTC. At 1300 UTC, it is seen that the measurements at 5 m and 10 m above ground show a shift to down-valley flow, while the readings from 15m and 20 m heights show that the flow is cross-valley. At 1400 UTC, we see the wind returns to down-valley direction and the thermal stratification is again built up (Figure 24). The wind speed measurements from Figure 31a, show that the westerly cross-valley flow is generally weaker in magnitude than the along-valley flow. In addition, it is observed that along-valley flow promotes velocity shear while cross-valley flow tends to destroy it. From the wind direction data in Figure 31b, we see that the wind shear is low due to the westerly cross-valley wind at 1030 UTC, in addition to this the thermal stratification is high (see Figure 30). This is the reason we see that the Ri number rise precipitously at 1030 UTC (Figure 24). Between 1030 UTC and 1200 UTC, we see both thermal stratification and high wind shear, resulting in Ri numbers close to 1 and sometimes below the critical value of 0.25. At 1200 UTC, the wind shear is again reduced due to the westerly cross-valley flow, while some thermal stratification is still present and thus the high Ri number is repeated. Therefore the sudden rise in the Ri number at 1030 UTC and 1200 UTC seen in Figure 24 could be explained by these cross-valley intrusions that occur in the valley wind system.

5.5 COAMPS simulation

Comparing COAMPS model results with lidar observations (Figure 28) we see that, COAMPS is able to model the larger-scale wind directions reasonably well. The two-layered flow structure with the lower level down-valley flow and the upper level up-valley flow is well represented by COAMPS. It is seen that there are deviations in the wind speed magnitude between the COAMPS results and the lidar measurements in the bottom 500 m of the valley. It has been observed in that the flow is thermally driven and there are strong thermal gradients within 500 m above the valley floor. The deviations of the COAMPS results from the lidar observations could be due to the lack of vertical resolution in the model in this region.

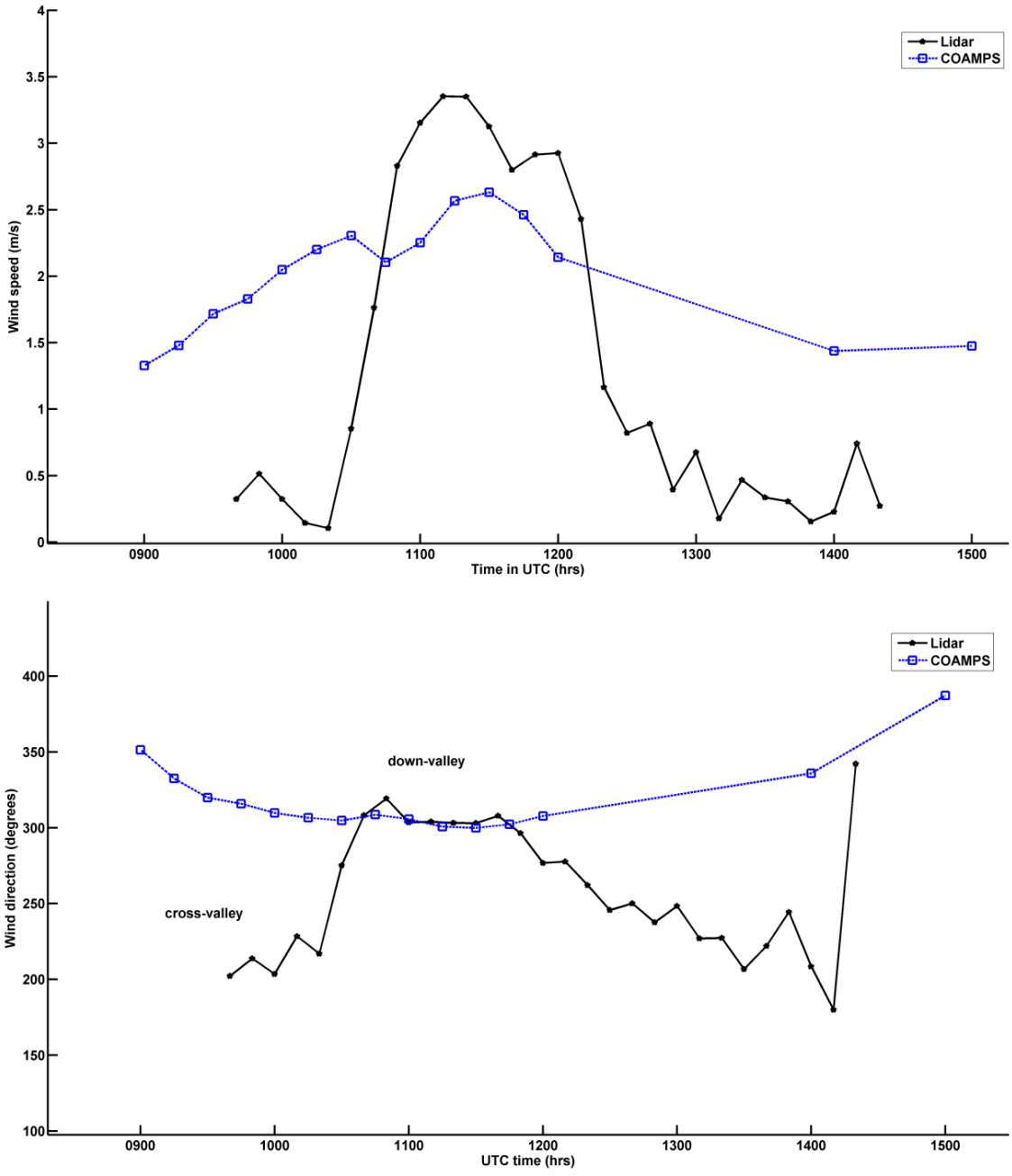


Figure 32. Comparison of a) wind speed and b) wind direction from the ASU lidar and COAMPS at 10 m height for the period under study.

A comparison of the wind speed and direction measured by the ASU lidar and predicted by COAMPS at 10 m height for the period under study is presented in Figure

32. It is seen that although COAMPS predicts the low level down-valley flow, it is unable to capture the periodic cross-valley intrusions detected by the ASU lidar (see Fig. 13). This might also be due to the lack of vertical resolution in the lower levels of the model. Increasing the number of vertical levels in this portion might help resolve the flow better. Doppler lidar measured flow patterns for such cases provide valuable opportunities to challenge high resolution mesoscale models.

5.6 Conclusions

A vector retrieval algorithm based on optimal interpolation and a high resolution COAMPS run is used to analyze the horizontal flow structure during the early morning period of March 27 2006, in the Owens Valley during the T-REX field experiment. It is observed from the lidar data that a two layer structure is formed within the valley starting at approximately 0930 UTC (0130 hr local time). Evidence presented in the figures shows that the lower level flow is down-valley, and presumably thermally driven, while the upper level is up-valley, synoptically forced and channeled due to the valley walls. The down-valley flow is observed to be quite weak with wind speeds in the range of 2-4 m s-1 and is limited to the bottom 200 m height.

A double layered cross-valley flow is observed by the ASU lidar at 1030 UTC. Observations from both the lidars indicate presence of westerly cross-valley intrusions occurring periodically. These westerly cross-valley intrusions are very close to the ground and at much lower magnitudes (0-1.5 m s^{-1}) compared to the along-valley flow. Based on the observations from the lidars and the tower measurements, it is observed that close to the ground, the flow alternates between being stratified and being well-mixed

vertically. The observations from the Doppler lidars are compared with the results from a high resolution COAMPS run, as well as, in-situ instrumentation present in the valley. It is seen that the lidar observations agree quite well with the measurements from in-situ instrumentation. However, it is observed that while COAMPS models the larger scale along-valley flows, it fails to capture some of the smaller scale features such as the periodic cross-valley intrusions observed by the ASU lidar. The present analysis showcases the ability of Doppler lidar to obtain spatially and temporally resolved vector fields in complex nocturnal valley flows. It is seen that the optimal interpolation technique is able to significantly preserve local spatial and temporal information captured by the lidar. It also provides evidence of the ability of sophisticated non-hydrostatic prediction systems, such as COAMPS, to predict the dominant large-scale flows over complex terrain with high degree of accuracy.

6. EFFECT OF TURBULENCE ON WIND POWER

6.1 Introduction

The recent surge in the technological development of wind turbines has seen development of turbines capable of producing power in the multi-MW range. With increased size and power capacity of wind turbines, come concerns of accurately predicting the power output from these turbines. Advances in wind turbine technology have resulted in wind turbines with hub-heights over 100 m and rotor diameters in excess of 80 m. Traditionally, wind measurements have been carried out using meteorological towers which provide point measurements at hub heights. With the increase in rotor swept area and hub heights, the impact of wind shear and turbulence intensity become increasingly relevant and point measurements from meteorological towers no longer are good representations of the wind interacting with the turbine. Remote sensing techniques – especially scanning coherent Doppler lidar, are able to rectify this by providing velocity measurements in three-dimensional space.

One of the main challenges of wind energy development is the variable nature of the wind resource. Uncharacteristic and extreme wind events have the potential to be disruptive to wind energy production and may require temporary shutdown of wind turbine operation to avoid structural damage. Therefore identification of such events and accounting their influence in wind resource assessment provides a more accurate estimate of the expected annual energy production from a particular wind farm site. Characterizing this variability as a function of space and time is essential to optimize wind energy output. In addition to the variability in wind speed and direction, the turbulence

parameters such as fluctuations of velocity components and direction fluctuations are important for determining the wind energy content.

The effect of wind speed turbulence and its effect on wind power production has been a subject of considerable research with Elliot and Cadogan (1990) showing the influence of wind shear and turbulence on wind turbine power curves. Wagner et al. (2009) showed that significant error in power measurement could result if the effect of wind shear was ignored and demonstrated the use of equivalent wind speed method for power curve measurement.

Wind speed turbulence is often accompanied by turbulence in wind direction. The power output is calculated assuming the wind interacts with the rotor parallel to its axis. Misalignment of the wind direction with the rotor axis can significantly reduce the power output. Akhmatov (2007) describes the influence of wind direction on wind power fluctuations. Pedersen (2004) using the 3D actuator disk model showed that the power output shows a $\cos^2$ relationship to the yaw error. But these studies do not capture the effect of fluctuations in the wind direction. As fluctuations in wind directions are expected along with fluctuations in the wind speed, their combined effect needs to be studied. In addition, it is not uncommon to expect shear in wind direction along with shear in wind speed with height on a wind farm. Therefore, the combined effects of the shear in wind direction and wind speed need to be considered for accurate assessment of the wind power output. In this chapter, theoretical formulations will be derived to account for the effects of directional turbulence and shear along with wind speed turbulence and shear and their effects on power output.

In section 6.2, we will derive an expression that accounts for turbulence in velocity and direction as well as shear in velocity and direction. An expression for equivalent wind speed will be developed. In section 6.3 we will use lidar data to test these formulations and quantify the effect of these on wind power production.

6.2 Theory

The effect of wind shear and wind direction on power output of wind turbines has been evaluated independently (Kaiser et al., 2003; Antoniou et al., 2009; Wagner et al., 2009 etc). These effects do not occur in isolation, but combined.

To start off, the power available in the wind can be expressed in the form of wind energy flux or kinetic energy flux (Wagner et al., 2009) defined as:

$$P(t) = \frac{1}{2}\rho C_P A U(t)^3 \tag{51}$$

where ρ is the density of air, C_P is the coefficient of power, A is the rotor swept area and $U(t)$ is the wind speed. To account for the possibility of azimuthal angle variations in the airflow we re-write the equation as:

$$P(t) = \frac{1}{2}\rho C_P A\left[U(t)\cos\alpha(t)\right]^3 \tag{52}$$

where $\alpha(t)$ is the angle of the wind to the turbine axis. To account for the influence of turbulence in wind speed and direction, we write:

$$U(t) = \overline{U} + u' \text{ and } \alpha(t) = \overline{\alpha} + \alpha' \tag{53}$$

where, $\overline{U}$ and $\bar{\alpha}$ are the 10 minute means of the wind speed and wind direction respectively. Substituting equation (53) into equation (52) we get:

$$\begin{aligned} P(t) &= \frac{1}{2}\rho C_P A[\overline{U}+u']^3[\cos(\overline{\alpha}+\alpha')]^3 \\ &= \frac{1}{2}\rho C_P A[\overline{U}+u']^3\left[1-\frac{(\overline{\alpha}+\alpha')^2}{2}\right]^3 \\ &= \frac{1}{2}\rho C_P A\overline{U}^3\left[1+\frac{u'}{\overline{U}}\right]^3\left[1-\frac{\overline{\alpha}^2}{2}\left(1+\frac{\alpha'}{\overline{\alpha}}\right)^2\right]^3 \\ &= \frac{1}{2}\rho C_P A\overline{U}^3\left[1+3\frac{u'}{\overline{U}}+\frac{6}{2!}\left(\frac{u'}{\overline{U}}\right)^2\right]\left[1-\frac{\overline{\alpha}^2}{2}\left(1+2\frac{\alpha'}{\overline{\alpha}}+\frac{\alpha'^2}{\overline{\alpha}^2}\right)\right]^3 \end{aligned} \tag{54}$$

The variation of the 10 minute mean power is considered most important for studying the wind farm characteristics. Therefore taking the 10 minute mean of the above equation:

$$\overline{P} = \frac{1}{2}\rho C_P A\overline{U}^3\left[1+3\left(\frac{\sigma_u}{\overline{U}}\right)^2\right]\left[1-\frac{\overline{\alpha}^2}{2}-\frac{\sigma_\alpha^2}{2}\right]^3 \tag{55}$$

where σ_u^2 is the variance of wind speed fluctuations and

σ_α^2 is the variance of direction fluctuations.

From equation (55) we can write the turbulent equivalent wind speed as:

$$U_T = \sqrt[3]{\overline{U}^3\left[1+3\left(\frac{\sigma_u}{\overline{U}}\right)^2\right]\left[1-\frac{\overline{\alpha}^2}{2}-\frac{\sigma_\alpha^2}{2}\right]^3} \tag{56}$$

The effect of wind shear can be accounted for by rewriting the wind speed in terms of the equivalent wind speed (see equation (4) in Wagner et al., 2009). Substituting the turbulent equivalent wind speed into the shear equivalent wind speed we get:

$$U_{eq} = \sqrt[3]{\frac{1}{A}\sum_{i=1}^{N} U_{Ti}^3 A_i} = \sqrt[3]{\frac{1}{A}\sum_{i=1}^{N} \overline{U}_i^3 \left[1+3\left(\frac{\sigma_{ui}}{\overline{U}_i}\right)^2\right]\left[1-\frac{\overline{\alpha}_i^2}{2}-\frac{\sigma_{\alpha i}^2}{2}\right]^3 A_i} \qquad (57)$$

where σ_{ui}^2 is the variance of velocity fluctuations at *i*-th level

α_i is the angle of the wind with respect to the rotor axis at *i*-th level

$\overline{U}_i$ is the wind speed at the *i*-th level

$\sigma_{\alpha i}^2$ is the direction fluctuations at the *i*-th level

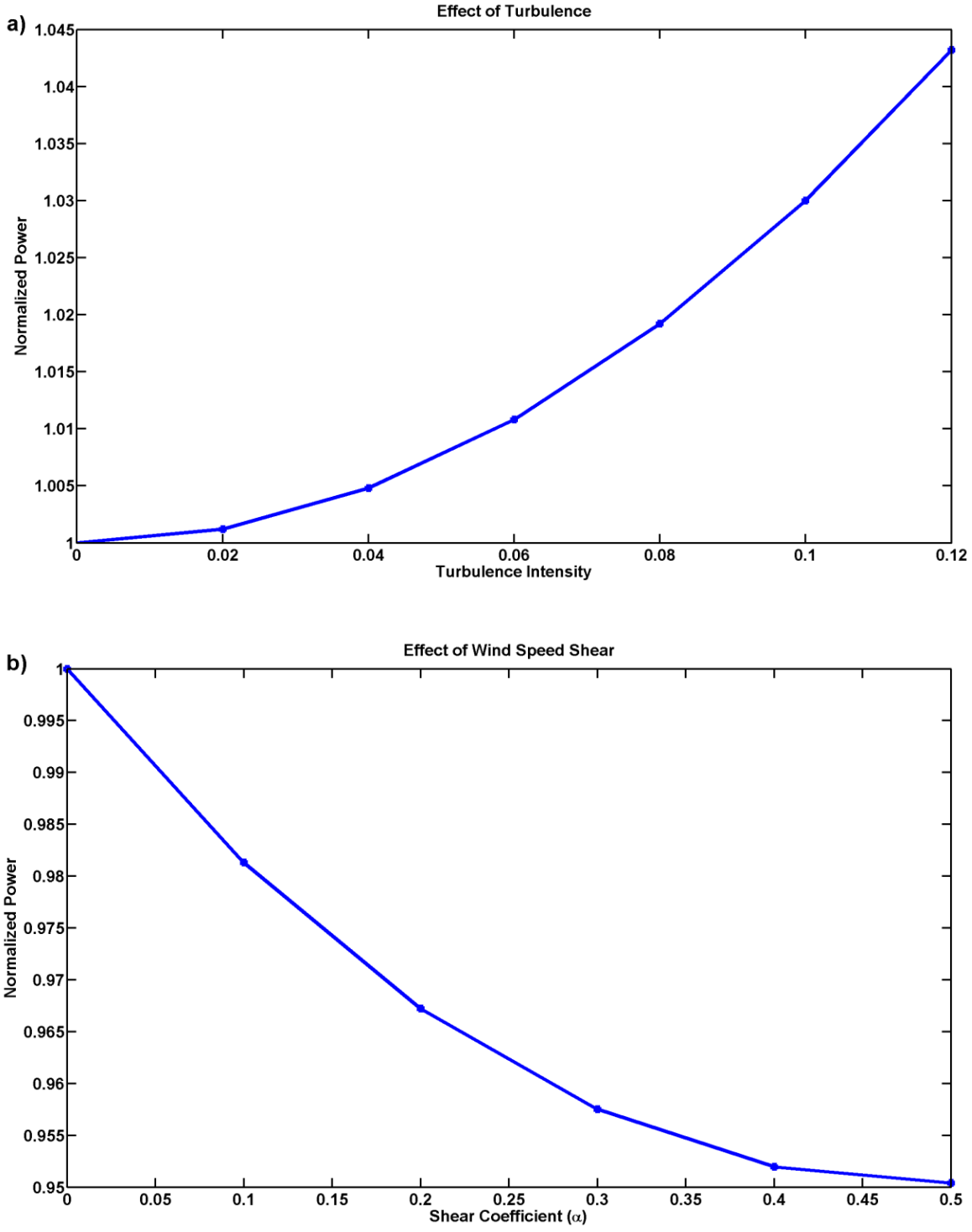
Therefore the equivalent wind energy content is given by:

$$\overline{P}_{eq} = \frac{1}{2}\rho C_P \sum_{i=1}^{N} \overline{U}_i^3 \left(1+3\left(\frac{\sigma_{ui}}{\overline{U}_i}\right)^2\right)\left[1-\frac{\overline{\alpha}_i^2}{2}-\frac{\sigma_{\alpha i}^2}{2}\right]^3 A_i \qquad (58)$$

6.3 Quantifying the Impact of Turbulence on Wind Power

From the equations derived in Section 6.2, it is possible to quantify the effect of turbulence and direction fluctuations on wind energy content. The wind energy content is strongly correlated with the wind power produced by the turbines and hence gives a reasonable estimate of the expected turbine performance in those conditions. Based on typical values of wind and direction shear as well as turbulence intensity, the variations in

wind energy content are estimated using the formulation in the previous section. These variations are presented in Figure 33.



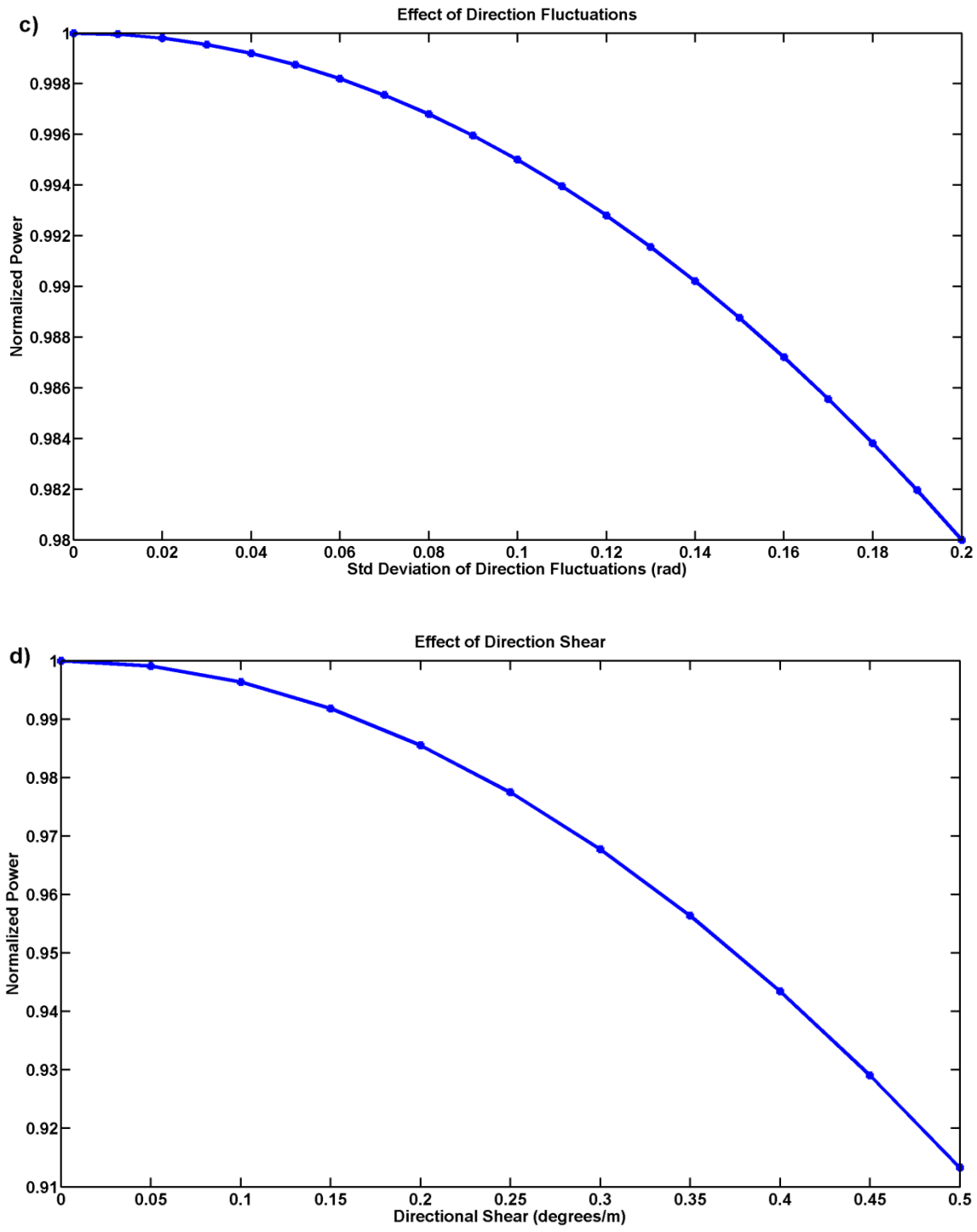
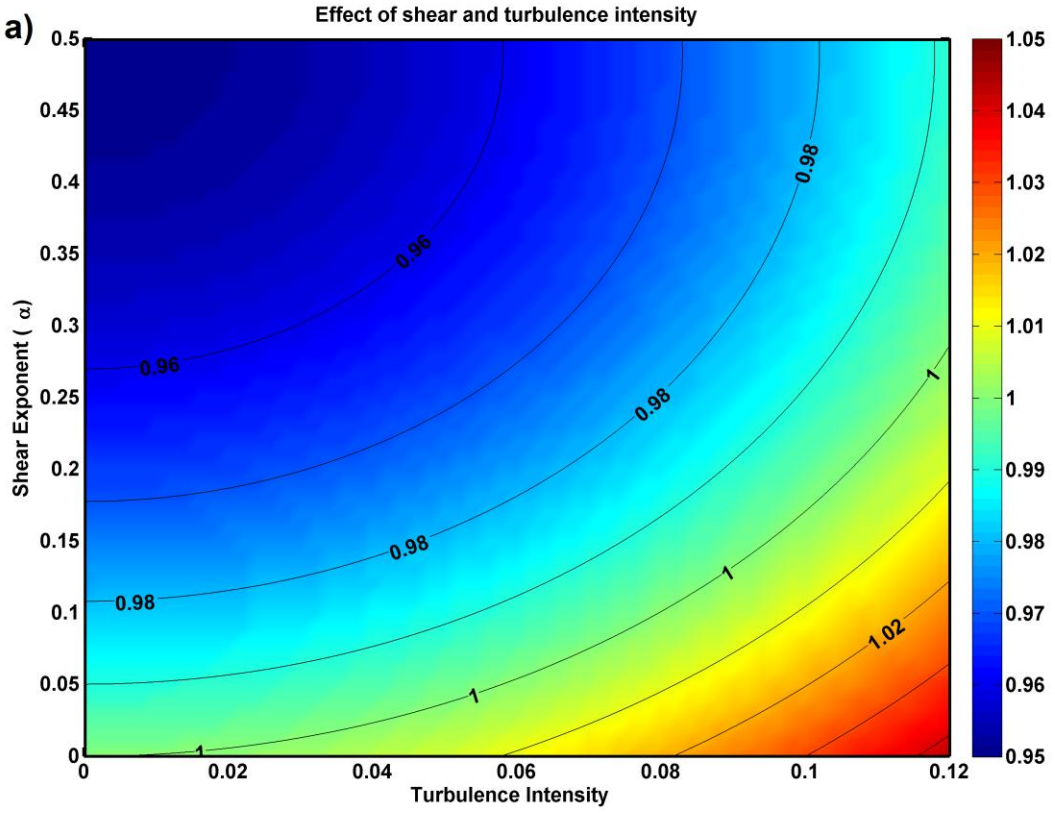


Figure 33. Effect of turbulence on wind energy content. (a) Impact of turbulence intensity. (b) Impact of wind speed shear. (c) Impact of direction fluctuations. (d) Impact of directional shear.

As seen from Figure 33(a), higher values of turbulence intensity increase the energy content of the wind. Various studies (for example: Akhmotov 2007; Tindal et al., 2008) have shown that turbulence intensity increases power production at lower wind speeds and lowers power production near the rated wind speeds. It is observed that wind speed shear, direction fluctuations and directional shear all reduce the available wind energy content.

Since, these four effects do not occur individually, but occur in combination or all together, studying the combined effects of these parameters is essential to understand the variability in available wind energy. In Figure 34, the combined effects of these turbulence parameters taken two at a time are shown.



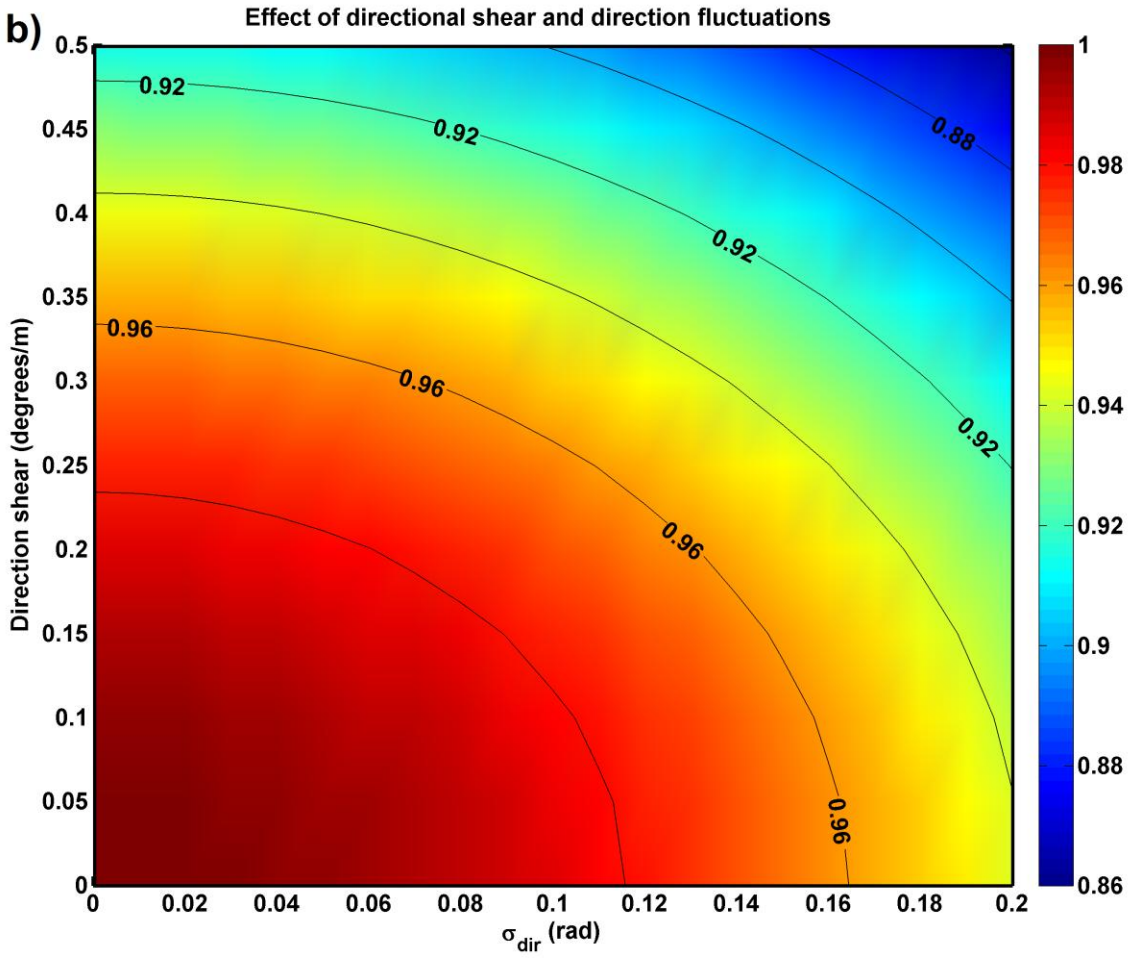


Figure 34. Combined effects of turbulence parameters on available wind energy content. The wind energy content is normalized with respect to the mean value in the absence of turbulence or shear. (a) Combined effect of wind speed shear and turbulence intensity. (b) Combined effect of directional shear and direction fluctuations.

The maps shown in Figure 34 show the variation of wind energy content due to the combined effects of various turbulence parameters. Such maps can be an important tool in wind turbine control to make the optimal control decision which will result in the maximum recovery of wind energy content. The availability of this information

combined with lidar's capability to make measurements several kilometers upstream of the wind turbine location will enable a planned control response.

6.4 Lidar Retrievals

The modified optimal interpolation (OI) technique (see Choukulkar et al., 2012b; Xu et al., 2007) is used to retrieve the two-dimensional velocity vectors from lidar data. The vectors retrieved through this technique will be used to estimate the turbulence intensity and variance of wind direction fluctuations. In order to validate the accuracy of the retrieval algorithm in estimating the turbulence parameters, comparisons with in-situ tower measurements will be made.

The data being used in this study is from a wind resource assessment site in USA. The authors are not at liberty to disclose the exact location and details of the site. The time period considered for this study is from June 20th to June 28th. The two dimensional velocity vectors were retrieved from the lidar data using the OI technique and the results were compared with an anemometer tower. Figure 35 shows the comparison of wind speed and wind direction measurements retrieved from a coherent Doppler lidar to that measured by an anemometer tower. The root mean square difference between the lidar and tower measurements is 1.47 m s^{-1} for the wind speed and 18^{o} for wind direction.

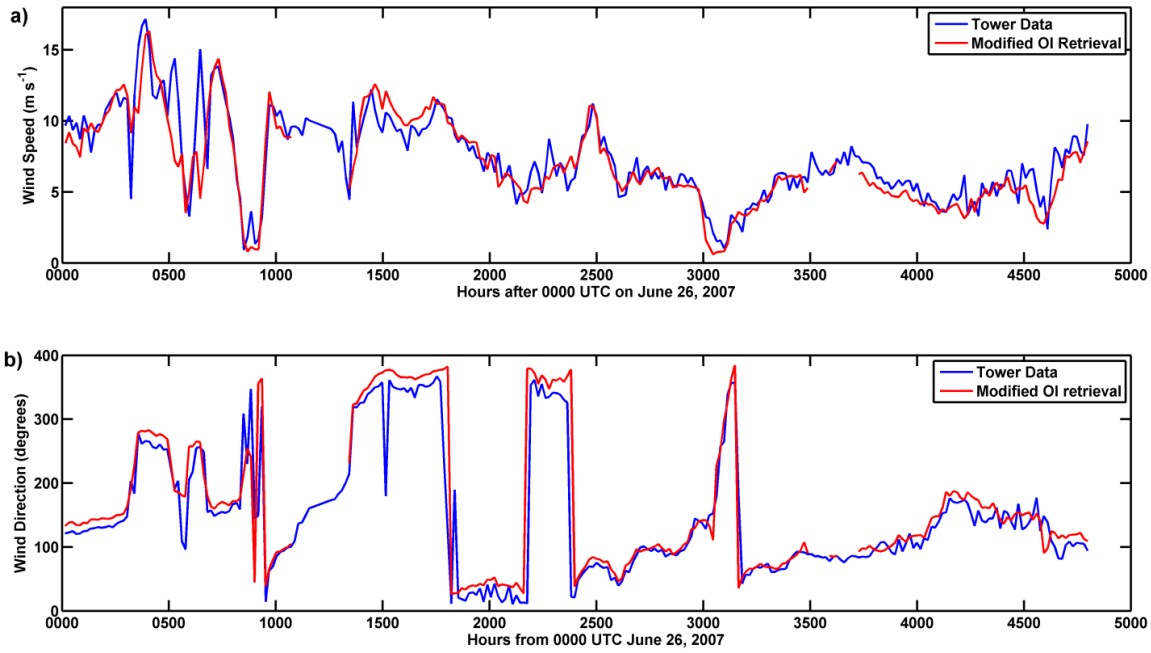


Figure 35. Comparison of lidar measurement (red) with tower measurements (blue) during the period of June 26th 2007 to June 27th 2007.

The data gathered at this wind development site featured several uncharacteristic and extreme wind events which could be potentially harmful for the structural integrity of wind turbines. Identifying and preparing for such events is essential to ensure safe wind farm operation and prolong the operational life of wind turbines. Some of the extreme wind events that were identified are presented here.

On June 26th at 0510 hrs local time, a sharp wind direction shear with height was observed. The lidar observations show the higher level (300 m above the ground) winds flowing from east to the west while the winds at lower heights (~50 m above ground) flowing from south-west to north-east. The vector retrievals on the lidar measurements made using the modified OI technique are presented in Figure 36. As observed from Figure 36, the wind direction changes by about 90 degrees within the turbine rotor swept

area. Identifying onset of such events and taking appropriate control action is essential to prevent high stresses in large wind turbine blades.

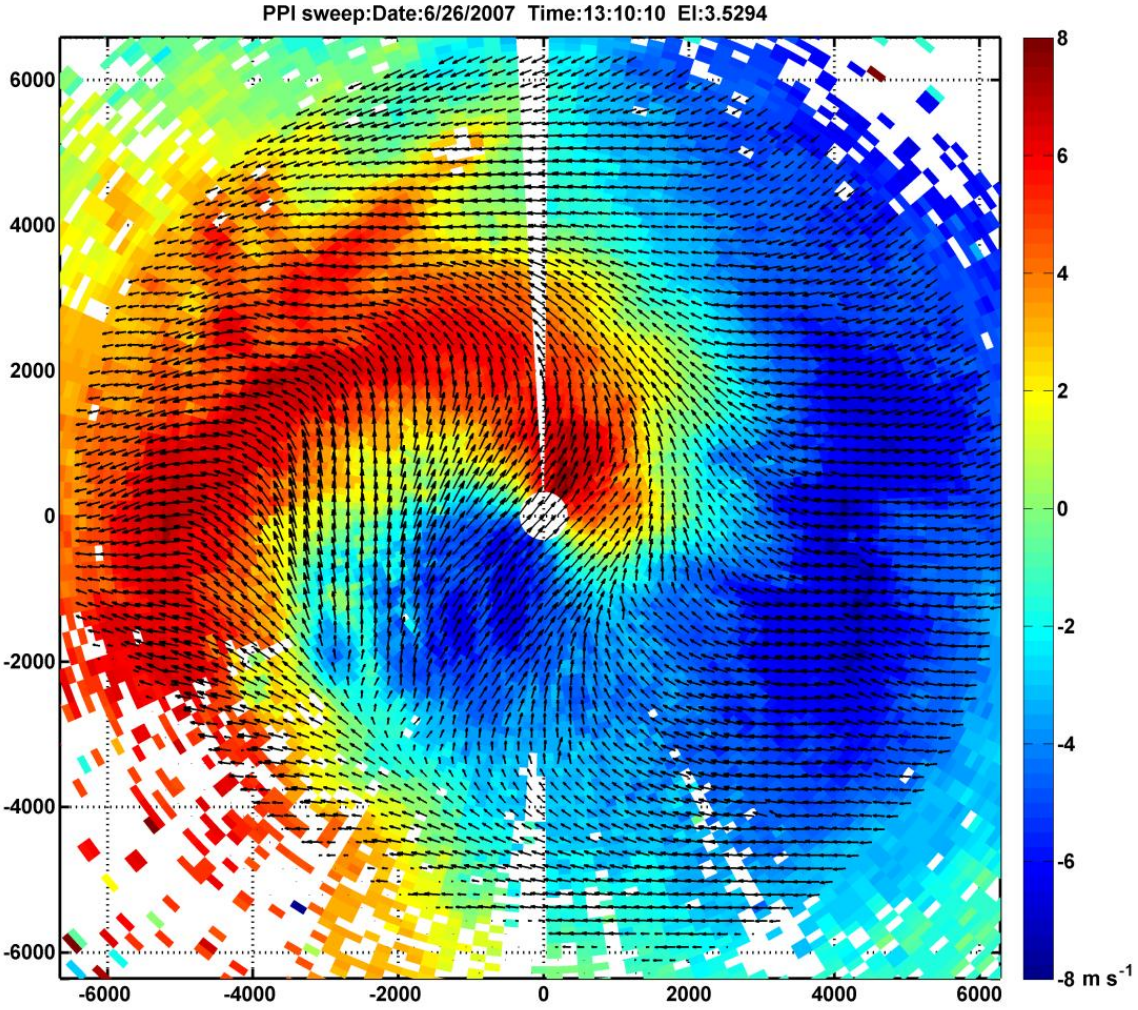


Figure 36. Vectors retrieved using the modified OI technique showing the sharp directional shear observed on June 26th, 0510 hrs local time on a PPI scan at 3.5 degree elevation. The colors indicate magnitude of radial velocity. Blue colors indicate flow towards the lidar and red color indicates flow away from the lidar.

On the same day, June 26th at 0600 hrs local time, lidar measurements captured gradual formation and decay of wind speed shear and wind directional shear. The

vertical structure of the wind field retrieved using the modified OI technique is presented in Figure 37. This plot was created by plotting the vectors obtained from all the elevation scans over a horizontal grid box of size 150 m x 150 m.

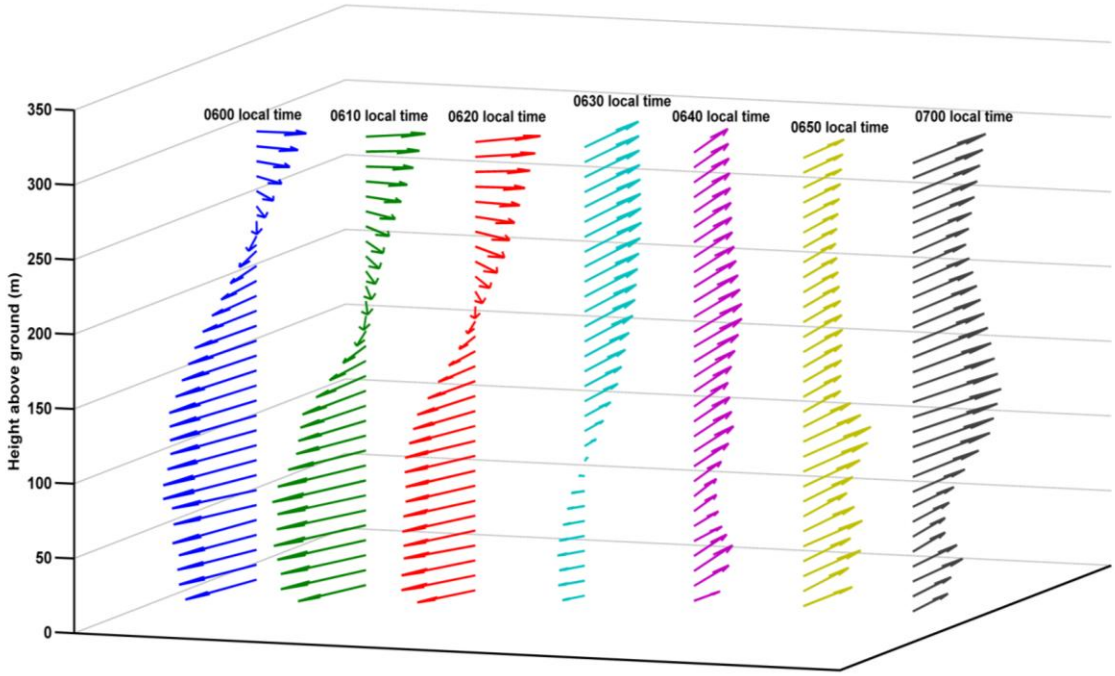


Figure 37. Evolution and decay of wind speed shear and wind direction shear over an hour as observed by the lidar on June 26th.

As seen in Figure 37, the wind starts to turn at height of 300 m around 0600 hrs local time. This direction shear is propagated downwards until at 0630 hrs local time we see the wind shear of 180 degrees between the heights of 50 m to 150 m, which falls within the rotor swept area of most modern wind turbines. At 0640 hrs local time, the direction shear is replaced by wind speed shear within the rotor swept area.

On June 27th, 2334 hrs local time, highly turbulent and converging flow was observed. The wind blowing from the west and south-east converged to flow north-east

as shown in Figure 38. Identification of such events is made possible with the help of remote sensing techniques as they can make measurements over large spatial extents.

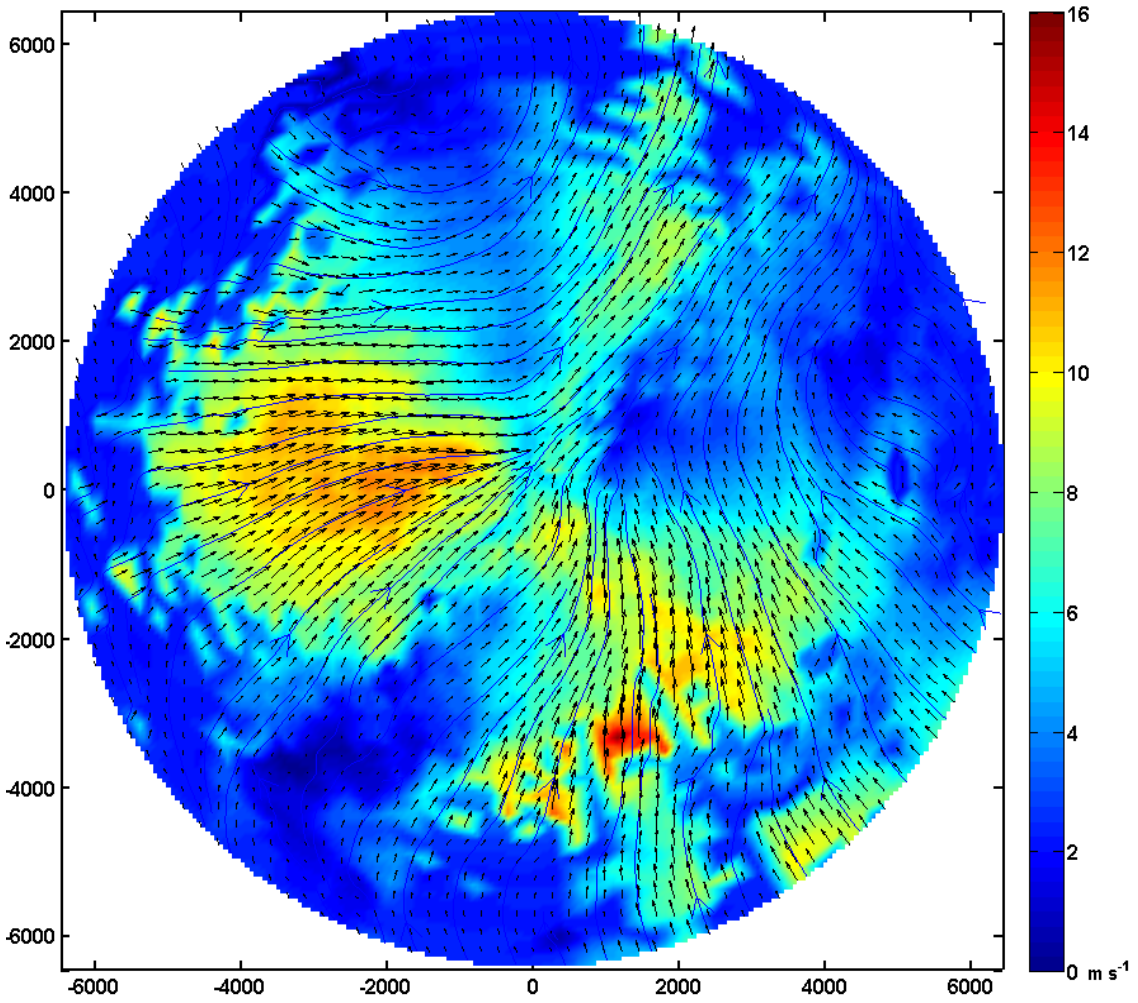


Figure 38. Lidar measurements showing highly turbulent and converging winds from the west and south-east on a 2 degree elevation PPI. Vectors retrieved using the modified OI technique. The colors indicate magnitude of wind speed.

Another feature of the boundary layer is the low level jet (see Figure 39). The low level jet is a layer of high speed flow formed during stably stratified conditions when the higher level flow is decoupled from the ground. Identification of the height, intensity

and frequency of occurrence of the low level jet is important for wind resource assessment as these events represent opportunities for higher wind energy production. Current mesoscale models lack the ability to correctly predict the correct height and intensity of low-level jets (Banta 2008) and hence, remote sensing measurements can reduce the uncertainty in detecting these events.

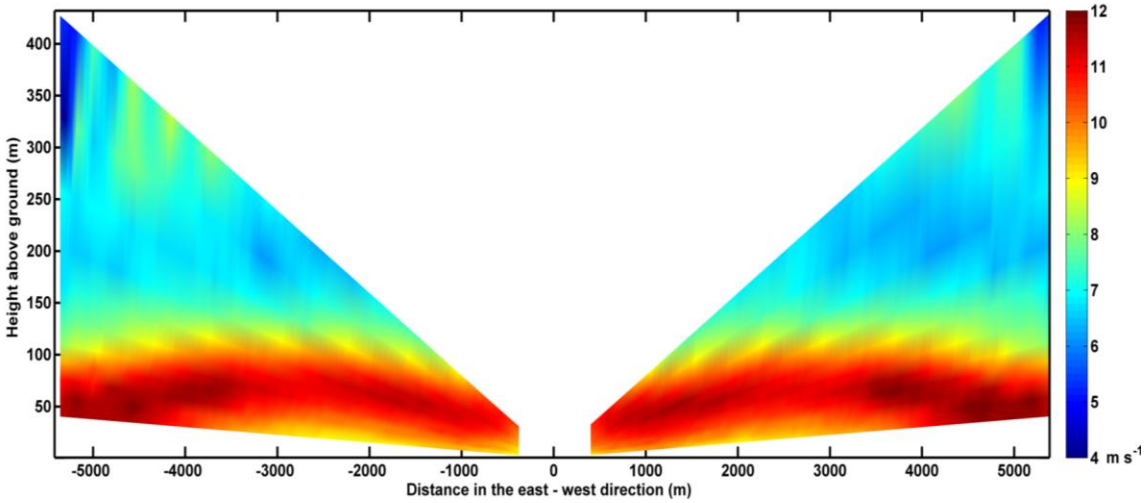


Figure 39. The low-level jet as measured by the coherent Doppler lidar on June 27th at 0040 local time. The colors indicate magnitude of wind speed. The vertical slices were created by collating data from a stack of PPI scans along a fixed azimuth.

With the help of the modified optimal interpolation technique, the turbulence parameters can be retrieved as a function of space and time. The comparisons of the lidar retrieved turbulence parameters with tower measurements are shown in Figure 40.

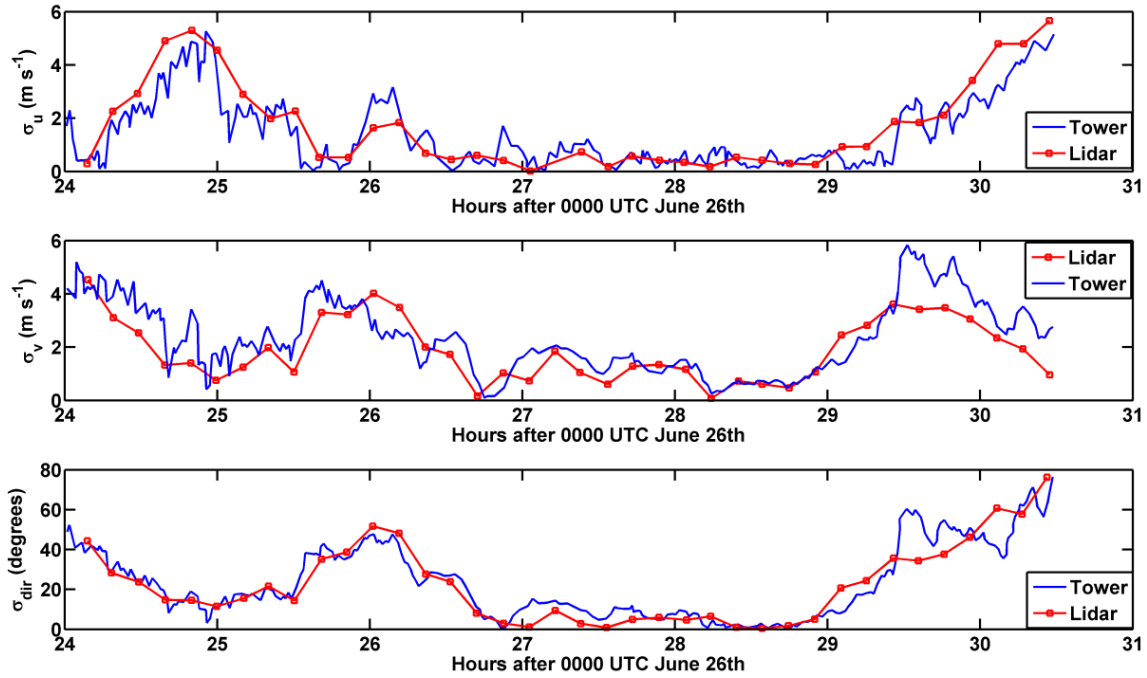


Figure 40. Comparison of turbulence parameters retrieved from lidar data to tower measurements. (a) Fluctuations in u-component measured by lidar and tower. (b) Fluctuations in v-component measured by lidar and tower. (c) Fluctuations in direction measured by lidar and tower.

Therefore the coherent Doppler lidar can play an important role in characterizing the wind resource in presence of turbulence by making high resolution measurements of turbulence parameters over the wind farm domain. Furthermore, by capturing local flow features as well as vertical structure of the flow field, Doppler lidar informed wind resource assessment can help make better decisions on wind turbine selection and siting.

6.5 Conclusion and Future Work

An expression for equivalent wind power has been derived which captures the effect of turbulence on available wind energy content. Through this equation the

available wind power can be estimated more accurately. Using this equivalent wind power equation, the effect of wind shear, turbulence intensity, direction shear and direction fluctuations on wind power are estimated. In addition, it is shown that this equation allows estimation of the combined effect of various turbulence parameters. This can be used to develop optimal control strategies that can deploy the control action that produces the maximum power in the given conditions.

The results from a field study at a wind development site show presence of extreme events such high shear flows, turbulent flows, low-level jets and directional shear with height. The equivalent wind speed and its impact on the kinetic energy content of the flow in these conditions are being studied. The wind power estimations based on these calculations will be compared to the actual wind power produced by the turbine. This will be done using data available from the Atlantic City wind farm site.

7. CONCLUSION AND FUTURE SCOPE

The mathematical formulation and validation of a modified optimal interpolation technique was demonstrated. In an attempt to develop a vector retrieval method that preserves the temporal and spatial resolution of the Doppler lidar while being computationally inexpensive, the Optimal Interpolation (OI) technique was adapted to the lidar remote sensing platform. Optimal interpolation is a statistical data assimilation technique that uses Bayesian probability theory to arrive at the maximum likelihood estimate of the true atmospheric state.

The Bayesian approach provides a unique framework that accounts for the occurrence of error in the observation data. This is well suited for lidar measurements which are characterized by random error due to shot noise and SNR (Signal to Noise Ratio) degradation with range. This method resulted in better preservation of local information compared to traditional least squares approaches while being less expensive computationally compared to variational methods (3DVAR and 4DVAR).

A lidar simulator was created to understand the error of representativeness and the vector retrieval error of the Doppler lidar. Results from a Large Eddy Simulation (LES) were used to recreate atmospheric flow conditions and lidar measurements were simulated within this field. Using the lidar simulator, the errors in lidar measurement process can be understood as a function of scanning configuration, look angles, turbulence intensity and scales of motion in the atmosphere. This study showed the ability of the modified OI technique to account for some errors of representativeness in

the lidar measurement. In addition, the effect of look angle on vector retrieval accuracy was examined.

This technique was applied to study a complex nocturnal event in the Owens Valley, California. Flow in mountain valleys is characterized by motions generated through interaction with the terrain and mountain slopes. These motions span a wide range of scales and are important in variety of applications such as dispersion modeling, boundary layer studies and wind farm characterization in complex terrain. Lidar based observations of these flows offers valuable insights in these studies due to the ability to obtain observations with high enough spatial and temporal resolution to capture the transient fluid dynamical features.

Choukulkar et al. (2012a) performed a study of wind events over the Owens Valley, California during the Terrain-induced Rotor Experiment (TREX). Using the OI technique, complex and transient flows, such as pulses of cross-valley momentum and their effect on along-valley winds, were identified. Results from a high resolution Couples Ocean/Atmospheric Mesoscale Prediction System (COAMPS) and other in-situ instruments were used to corroborate and complement the lidar observations.

The coherent Doppler lidar is emerging as a potentially transformative tool in the field of wind energy resource assessment. The ability of the coherent Doppler lidar to make measurements with high spatial and temporal resolution over wide spatial extents lends itself to improved wind resource assessment as has been shown through numerous studies (Krishnamurthy et al., 2012; Hannon et al., 2008; and Smith et al., 2006). Lidar remote sensing can be effectively used for identification of fluid dynamic processes

occurring over a wide range of scales. This leads to accurate characterization of wind turbine inflow parameters and allows better understanding of how different atmospheric conditions and wind events will influence wind power output. In addition, an accurate characterization of the boundary layer at the wind farm site is required to accurately characterize local features such as the low-level jet, density gradients which play an important role in wind farm power output.

An equation for the equivalent wind power in the presence of turbulence intensity, wind shear, direction fluctuations and directional shear was developed. The effect of these parameters on wind energy content was demonstrated. It was shown that through this equation, the combined effect of various turbulence parameters can be mapped, leading to development of better control strategies.

This work aims to establish a foundation for lidar data analysis for various applications such as boundary layer studies, mountain meteorology, dispersion modeling and wind energy resource assessment. Future work will enable application of these techniques to analyzing wind turbine wakes, ramp detection and short term forecasting, studying boundary layer processes such as turbulence over plant canopies and interaction between large scale above canopy flow with the Reynolds stress events within the canopy and important lower atmospheric phenomena such as the Madden-Julian Oscillation. Some of the studies that will be undertaken in the future are explained in the following sections.

7.1 Investigation of Wind Turbine Wakes

Advances in wind turbine technology have resulted in wind turbines with hub-heights over 100 m and rotor diameters in excess of 80 m. As a result, wake regions of these wind turbines have considerable effect on fatigue loading and power output of downstream wind turbines. Studies have shown that downstream turbines experience an increase in fatigue loading by 20% on an average even when separated by distance of 8D, where D is the turbine rotor diameter (Crespo et al, 1999), while power production can reduce by 40% (Mechali et al., 2006). Therefore accurate wake modeling is central to efficient wind farm design.

Fluid motions within a wake span a wide range of scales. The velocity deficit and turbulence intensities within a wake region are a complex function of turbine inflow conditions, ambient turbulence intensity, atmospheric stability and topographic effects. In addition wind variations on scales larger than wake diameter induce wake meandering which is a dominant source of error in wake modeling (Crespo et al., 1999). Ground based scanning Doppler LIDAR (Light Detection and Ranging) remote sensing can simultaneously capture the spatial and temporal evolution of a wide range of scales important for characterizing wake dynamics and meander (Larsen et al. 2008).

Several studies have been conducted (both wind tunnel and field experiments) to characterize the behavior of the flow in the wake region (Vermeer et al. 2003). Wind tunnel experiments allow studying the impact of each parameter on wake behavior individually. However, reproduction of real atmospheric conditions is difficult in wind tunnels and therefore results differ when compared to observed values. For example,

Högström et al. (1988) report turbulence intensity measurements in wind tunnels to be higher by a factor of two, while Helmis et al. (1995) reported an overestimation of the length of the near wake region. Recent efforts have tried to recreate real atmospheric conditions in wind tunnels. Yang et al. (2011) simulated atmospheric boundary layer condition in their wind tunnel at Iowa State University to study wake characteristics of horizontal axis wind turbines. Chamorro and Porté-Agel (2010) simulated effects of thermal stability in a wind tunnel to study wake effects.

Field experiments allow measurement of wake characteristics of complex wind flow interacting with the turbine rotor. Remote sensing techniques have proven to be effective for wake measurements as they allow measurements at multiple locations without flow distortion. Barthelmie et al. (2003) performed SODAR (Sonic Detection and Ranging) measurements of wakes behind an offshore wind turbine. However, the measurement of the wake profile was limited to one fixed downstream distance. Bingöl et al. (2010) used a nacelle mounted continuous wave Doppler lidar to make wake measurements. Using several scanning configurations, line measurements of wake deficits at several downstream locations was performed. This technique is not suitable for measuring large scale meander, especially in the presence of yaw error. Käsler et al. (2010) performed wake deficit measurements of a multi-megawatt wind turbine using a pulsed Doppler lidar. Here, line of sight velocity was used to estimate the wake deficits. However, this methodology is limited to time periods when the wind direction is aligned with the line of sight and fails to measure deficits in presence of meander. Therefore, a methodology is required that can fully utilize the three-dimensional measurements made by a lidar and retrieve velocity and turbulence information for a wide range of motions.

The coherent Doppler lidar can produce data rich three dimensional measurements over a wind farm site allowing identification of fluid dynamic processes occurring over a wide range of scales (Pichugina 2010). This leads to accurate characterization of wind turbine inflow parameters and allows better understanding of how different atmospheric conditions and wind events influence wind power output. In addition, the coherent Doppler lidar allows measurement of the horizontal and vertical flow structure with high spatial and temporal resolution around wind turbines which can be effectively used to characterize velocity deficits and turbulence within wakes. Data collected using the Doppler lidar can be used to study the impact of inflow conditions, atmospheric stability and ambient turbulence on wake characteristics such as deficit, turbulence intensity within wakes and wake meander. This research will answer questions such as: (1) how does wind shear present in the inflow affect vertical structure of the wake? (2) How does stability impact wake diameter? (3) Does ambient turbulence intensity influence wake deficit magnitude? (4) How different inflow conditions affect wake meander? (5) Can wake meander be characterized using concepts from plume dispersion modeling?

Larsen et al. (2008) point out that effects of meander could result in: (1) velocity deficits being less than that computed and (2) turbulence intensities higher than that computed by conventional wake models. There are conflicting theories on the size of scales responsible for meander of wakes (Crespo et al., 1999). While some argue that meander is caused by scales of motion larger than the size of the wake (Baker et al., 1985; Taylor 1993), others argue that all scales of motion influence meander (Högström et al., 1988).

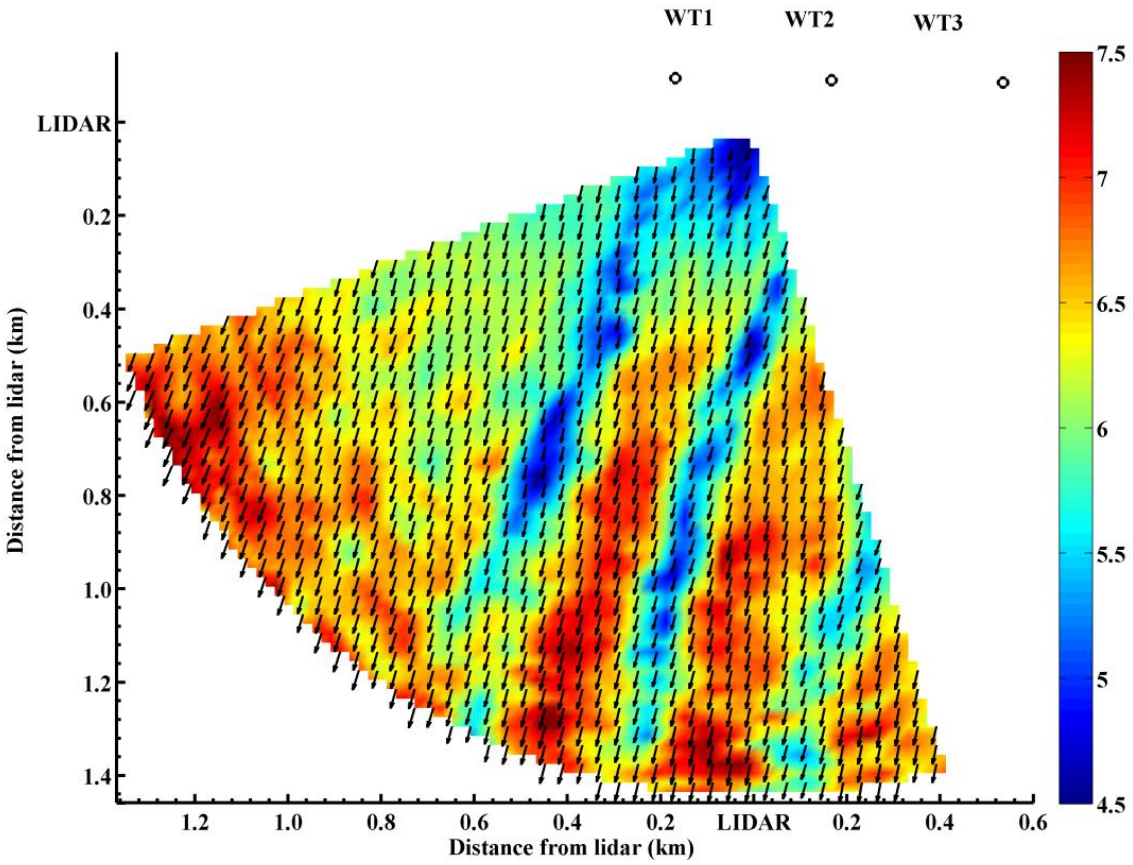


Figure 41. Vector retrieval of wake behind turbines (WT1, WT2, WT3) measured by Doppler lidar during Great Lakes Experiment using OI technique. Colors indicate magnitude of wind speed in m s^{-1}. See Barthelmie et al. (2012) for poster description of Part 1 of the experiment in which this data was acquired.

The approach used to model wake meander will be similar to that of Larsen et al. (2008), where the transverse and vertical transport of velocity deficit within the wake is modeled as if it acted as a passive tracer. However, Larsen et al. (2008) assume the wake to be rotationally symmetric and invoke Taylor's hypothesis for the downward advection of the deficits. These assumptions are reasonable when modeling large scale effects in flat terrain, however the validity of these assumptions are in question in presence of wind

shear and complex terrain effects. In addition, the Mann spectral model (Mann 1994) used in this method is applicable only for neutral stratification.

In the case of the Doppler lidar, the assumption of Taylor's hypothesis is limited to the time taken by the lidar to complete one scan. Lidars have already been successfully employed to measure plume and puff dispersion in the atmospheric boundary layer (Eberhard et al., 1987; Collier et al., 2005; Choukulkar et al., 2011). In this framework, the role of different scales can be split as follows: (1) scales equal to or less than the size of the wake are responsible for mechanical mixing and wake widening (relative dispersion) and (2) scales greater than the size of the wake are responsible for large scale motion of the wake (meander). A long term goal of this research is to develop parameterizations that account for within wake dynamics and large scale meander under different atmospheric forcings such as stability, wind shear and ambient turbulence.

7.2 Short-term Forecasting and Ramp Detection

One of the main drivers of high power generation costs from wind turbines is the low capacity factor due to "wind resource variations" (US Department of Energy, 2010) that cause unpredictability in power production. Since a significant portion of the cost per kilowatt hour of many wind power projects is tied to the negotiated rates for borrowed capital, characterizing the winds with increased precision (and thereby improving confidence intervals) should reduce risk for lending institutions and may favorably affect rates (US Department of Energy Market report 2010). Reducing the cost per kilowatt hour would sharpen the competiveness of this important renewable energy source.

Another important source of power variability in wind farms is the presence of turbulence and shear in the wind. The effect of wind speed turbulence and its effect on wind power production has been a subject of considerable research with Elliot and Cadogan (1990) showing the influence of wind shear and turbulence on wind turbine power curves. Wagner et al. (2009) showed that significant error in power measurement could result if the effect of wind shear was ignored and demonstrated the use of equivalent wind speed method for power curve measurement. The presence of turbulence and shear tend to modify the original power curve supplied by the manufacturers as has been demonstrated by Kaiser et al. (2007) and many others. In addition to turbulence, the atmospheric stability also plays a role in wind power (Sumner and Mason 2006; Wharton and Lundquist 2010; etc).

An unsolved challenge remains accurate prediction of wind ramp events that can change the power output from wind farms dramatically. Ramp events are sudden increase or decrease in wind speeds over a wind farm area which can lead to large power output changes in a short time (1 to 3 hours). Wind ramp events occur due to a combination of mesoscale and local meteorological effects. Mesoscale effects such as cold fronts, warm fronts, thunderstorm outflow, sea-breeze etc combined with local effects such as decoupling of the winds in the surface layer from ground, local radiation budget etc lead to upward ramp or downward ramp events. Therefore understanding and predicting wind ramp events requires knowledge of processes occurring over a wide range of scales. Attempts at forecasting ramp events range from statistical and probabilistic methods (see for example Cuttler and Kay 2007; Bossavy et al., 2010) to using numerical weather prediction (NWP) models (see Wharton et al., 2011; Storm et

al., 2009; etc). Both approaches have seen limited success with regards to prediction accuracy.

NWP model have shown considerable potential in accurately predicting the synoptic scale events that are precursors for most ramp events. Wharton et al. (2011) describe case studies done using WRF (Weather Research and Forecasting) model for several wind ramp events. It has been observed through several studies (e.g.: Storm et al., 2009; Gallus and Deppe 2011; Banta et al., 2010; etc) that NWP models exhibit high uncertainty in predicting local meteorological processes such as low-level jets, formation of stable boundary layer, density gradients etc which are local precursors to wind ramp events.

Sensitivity studies of NWP models have shown that these models are sensitive to certain geographical locations. Accurate surface layer measurements in these geographical "hot spots" may lead to improvement in prediction skills of the NWP models. In addition, frequency of assimilation of measurements may to play a role in forecast skill. Determining the optimum assimilation frequency so as to get maximum forecast skill for given computing constraints may lead to a possible solution for maximizing the wind energy potential for a site. Therefore a combination of NWP models with ground based remote sensing measurements have potential to provide a complete picture of how the synoptic scale processes downscale to the micro-meteorological scales. This research can then provide pathways on how surface-layer parameterization schemes in NWP models can be optimized to capture these effects.

BIBLIOGRAPHY

Allwine K.J., Flaherty J.E., (2006): Joint Urban 2003 - Study overview and instrument locations. *Pacific Northwest National Laboratory Technical Report PNNL-15967*, Richland: Washington, DC, 92.

Akhmatov V. (2007): Influence of wind direction on intense power fluctuations in large offshore wind farms in the North Sea. *Wind Engineering*, **31**, 1, 59-64.

Antoniou I., Pedersen S.M., Enevoldsen P.B. (2009): Wind shear and uncertainty in power curve measurement and wind resources. *Wind Engineering*, **33**, 5, 449-468.

Baker, E.H, (1992): Design of the Navy's multivariate optimum interpolation analysis system. *Weather and Forecasting*, 7, 220-231.

Banakh V., Smalikho I. (1994): Estimation of the turbulence energy dissipation rate from the pulsed Doppler lidar data. *Journal of Atmospheric and Oceanic Optics*, **10**, 12, 957-965.

Banta R.M., Newsom R.K., Lundquist J.K., Pichugina Y.L., Coulter R.L., Mahrt L. (2002): Nocturnal low level jet characteristics over Kansas during CASES-99. *Boundary-Layer Meteorology*, **105**, 221-252.

Banta R.M., Darby L.S., Kaufmann P., Levinson D.H., Zhu C-J. (1999): Wind-flow patterns in Grand Canyon as revealed by Doppler lidar. *Journal of Applied Meteorology*, **38**, 1069-1083.

Banta R.M., Pichugina Y.L., Kelley N.D., Jonkman B., Brewer W.A. (2008): Doppler lidar measurements of the Great Plains low-level jet: Applications to wind energy. *14th International Symposium for the Advancement of Boundary Layer Remote Sensing*.

Bergthorsson P., Doos B. (1955): Numerical weather map analysis. *Tellus*, **7**, 329-340.

Browning K.A., Wexler R. (1968): The determination of kinematic properties of a wind field using Doppler radar. *Journal of Applied Meteorology*, **7**, 105–113.

Bouttier F., Courtier P. (2001): Data assimilation concepts and methods. *Meteorological Training Course and Lecture Series*. ECMWF, March 1999.

Calhoun, R., Heap R., Princevac M., Newsom R., Fernando H., Ligon D. (2006): Virtual towers using coherent Doppler lidar during the Joint Urban 2003 Experiment. *Journal of Applied Meteorology*, **45**, 1116-1126.

Choukulkar A., Calhoun R., Fernando H.J.S. (2011): The use of lidar-detected smoke puff evolution for dispersion calculations. *Meteorological Applications*, **18** (2), 188-197.

Choukulkar A., Calhoun R., Billings B.J., Doyle J.D. (2012a): Investigation of a complex nocturnal flow in Owens Valley using lidar and COAMPS. *Boundary Layer Meteorology* (Accepted)

Choukulkar A., Calhoun R., Billings B.J., Doyle J.D. (2012b): A modified optimal interpolation technique for vector retrieval for coherent Doppler lidar. *IEEE Geoscience and Remote Sensing Letters* (Accepted).

Collier C.G., Davies F., Bozier K.E., Holt A.R., Middleton D.R., Pearson G.N., Siemen S., Willetts D.V., Upton G.J., Young R.I. (2005): Dual-Doppler lidar measurement for improving dispersion models. *Bulletin of the American Meteorological Society*, **86**, 825–838.

Courtier P., Andersson E., Heckley W., Pailleux J., Vasiljevic D., Hamrud M., Hollingsworth A., Rabier F., Fisher M. (1998): The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part 1: formulation. Quarterly Journal of the Royal Meteorological Society, **124**, 1783-1807.

Dabas, A. M., P. Drobinski, and P. H. Flamant, 1998: Chirp induced bias in velocity measurements by a coherent Doppler CO2 lidar. J. Atmos. Oceanic Technol., 15, 407–415.

Daley R. (1991): Atmospheric Data Analysis. *Cambridge University Press*, pp 457.

Doyle J.D., Durran, D.R. (2002): The Dynamics of Mountain-Wave Induced Rotors. *Journal of Atmospheric Sciences*, **59**, 186-201.

Dunne F., Simley E., Pao L.Y. (2011): Lidar wind speed measurement analysis and feed-forward blade pitch control for load mitigation in wind turbines. *NREL Subcontract Report*, NREL/SR-5000-52098.

Elliot D.L., Cadogan J.B. (1990): Effects of wind shear and turbulence on wind turbine power curves, *European Wind Energy Conference 1990*, September 10–14, Madrid, Spain.

Frehlich R., 2004: Velocity error for coherent Doppler lidar with pulse accumulation. J. Atmos. Oceanic Technol., 21, 905–920.

Gal-chen T., Xu M., Eberhard W. (1992): Estimations of atmospheric boundary layer fluxes and other turbulence parameters from Doppler lidar data. *Journal of Geophysical Research*, **97**, 409-418.

Grubišić V., et al (2008): The Terrain-Induced Rotor Experiment: An overview of the field campaign and some highlight of special observations. *Bulletin of the American Meteorological Society*, **89**, 1513-1533.

Grubišić V., Billings B.J. (2007): The Intense Lee-Wave Rotor Event of Sierra Rotors IOP8. Journal of Atmospheric Sciences, **64**, 4178-4201.

Gandin, L. (1963): Objective analysis of the meteorological fields (Lenningrad: Gridromet). English translation (Jerusalem: Israel Program for Scientific Translation).

Hannon S.M., Barr K., Novotny J., Bass J., Oliver A., Anderson M. (2008): Large scale wind resource mapping using a state-of-the-art 3-D scanning lidar. *European Wind Energy Conference*, 2008.

Hill M., Calhoun R., Fernando H.J.S., Wieser A., Dornbrack A., Weissmann M., Mayr G., Newsom R. (2010): Coplanar Doppler lidar retrievals of rotors from T-REX. *Journal of the Atmospheric Sciences*, **67**, 713-729.

Hollingsworth, A., Lonnberg P. (1986): The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38**(A), 111-136.

Hodur, R.M. (1997): The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Monthly Weather Rev*iew, **125**, 1414-1430.

Jiang Q., Doyle J.D. (2008): Diurnal Variation of Downslope Winds in Owens Valley during the Sierra Rotor Experiment. *Monthly Weather Review*, **136**, 3760-3780.

Lenschow D., V. Wulfmeyer, and C. Senff, 2000: Measuring second through fourth-order moments in noisy data. J. Atmos. Oceanic Technol., 17, 1330–1347.

Lin, C., Xia Q., Calhoun R. (2008): Retrieval of urban boundary layer structures from Doppler lidar data. Part II: Proper orthogonal decomposition. *Journal of Atmospheric Sciences*, **65**, 21-42.

Lorenc A. (1995): Atmospheric Data Assimilation. *Second WMO Symposium on Assimilation of Observations in Meteorology and Oceanography*, Tokyo, Japan, March 1995.

Lorenc, A. (1981): A Global three dimensional multivariate statistical analysis system. *Monthly Weather Review*, **109**, 701-721.

Kahn B.H., Chan W., Lester P.F. (1997): An investigation of rotor flow using DFDR data. *Proceedings of the Seventh Conference on Aviation, Range, and Aerospace Meteorology*, Long Beach, CA, American Meteorological Society, 206–210.

Kaiser K., Langreder W., Hohlen H., Højstrup J. (2007): Turbulence correction for power curves. *Wind Energy*, DOI: 10.1007/978-3-540-33866-6_28, 159-162.

Kongara S., Calhoun R., Choukulkar A., Boldi M.O. (2012): Velocity retrievals for coherent Doppler lidar. *International Journal of Remote Sensing* (In Press).

Krishnamurthy R., Choukulkar A., Calhoun R., Fine J., Oliver A., Barr K.S. (2012): Coherent Doppler lidar for wind farm characterization. *Wind Energy*, DOI: 10.1002/we.539.

Newsom R., Calhoun R., Ligon D., Allwine J. (2008): Linearly Organized Turbulence Structures Observed Over a Suburban Area by Dual-Doppler Lidar. *Boundary Layer Meteorology*, **127** (1), 111-130.

Newsom R., Chai T., Lin C. (2004): Retrieval of Microscale Flow Structures from High Resolution Doppler Lidar Data Using an Adjoint Model. *Journal of Atmospheric Sciences*, **61**, 1500-1520.

Pedersen T.F. (2004): On wind turbine power performance measurements at inclined flow. *Wind Energy*, **7**, 163-176.

Porch W.M., Clements W.E., Coulter R.L. (1991): Nighttime valley waves. *Journal of Applied Meteorology*, **30**, 145–156.

Princevac M., Hunt J.C.R., Fernando H.J.S. (2008): Quasi-Steady Katabatic Winds on Slopes in Wide Valleys: Hydraulic Theory and Observations. *Journal of Atmospheric Sciences*, **65**, 627-643.

Panofsky H. (1949): Objective weather-map analysis. Journal of Applied Meteorology, **6**, 386-392.

Parrish D.F., Derber J.C. (1992): The national meteorological center's spectral statistical interpolation analysis system. *Monthly Weather Review*, **120**, 1747-1763.

Rye, B. J. and Hardesty, R. M.: 1993, 'Discrete Spectral Peak Estimation in Incoherent Backscatter Heterodyne Lidar. I. Spectral Accumulation and the CramerRao Lower Bound', IEEE Trans. Geo. Sci. Remote Sensing 31, 16–27.

Schmidli J., Poulos G.S., Daniels M.H., Chow F.K. (2009): External Influences on Nocturnal Thermally Driven Flows in a Deep Valley. *Journal of Applied Meteorology Climatology*, **48**, 3-23.

Smith D.A., Harris M., Coffey A.S., Mikkelsen T., Jørgensen H.E., Mann J., Danielian R. (2006): Wind lidar evaluation at the Danish wind test site in Høvsøre. *Wind Energy*, **9** (1-2), 87-93.

Wagner R., Antoniou I., Pedersen S.M., Courtney M.S., Jorgensen H.E. (2009): The influence of wind speed profile in wind turbine performance measurements. *Wind Energy*, **12**, 348-362.

Weissmann M., Dornbrack A., Doyle J.D. (2009): Vorticity from line-of-sight lidar velocity scans. *Journal of Atmospheric and Oceanic Technology*, **26**, 2683-2690.

Whiteman C.D. (1990): Observations of Thermally Developed Wind Systems in Mountainous Terrain. Atmospheric Processes Over Complex Terrain, *Meteorological Monographs*, **23**, 45, American Meteorological Society, Boston, Massachusetts, 5-42.

Whiteman C.D. (2000): Mountain Meteorology: Fundamentals and Applications. *Oxford University Press*.

Whiteman C.D., Doran C.J. (1993): The Relationship between Overlying Synoptic Scale Flows and Winds within a Valley. *Journal of Applied Meteorology*, **32**, 1669-1682.

Whiteman C.D., Hoch S.W., Poulos G.S. (2009): Evening Temperature Rises on Valley Floors and Slopes: Their Causes and Their Relationship to the Thermally Driven Wind System. *Journal of Applied Meteorology and Climatology*, **48**, 776-788.

Xia Q., Lin C., Calhoun R., Newsom R. (2008): Retrieval of urban boundary layer structures from Doppler lidar data. Part I: Accuracy assessment. *Journal of Atmospheric Sciences*, **65**, 3-20.

Xu Q., Gong J. (2003): Background error covariance functions for Doppler radial-wind analysis. *Quarterly Journal of the Royal Meteorological Society*, **129**, 1703-1720.

Xu, Q., Liu S., Xue M. (2006): Background Error Covariance Functions for Velocity Wind Analysis using Doppler Radar Radial-Velocity Observations. *Quarterly Journal of the Royal Meteorological Society*, **132**, 2887-2904.

Xu, Q., Nai K., Wei L. (2007): An innovation method for estimating radar radial velocity observation error and background wind error covariances. *Quarterly Journal of the Royal Meteorological Society*, **133**, 407-415.

APPENDIX – A

DERIVATION OF THE OPTIMAL INTERPOLATION COST FUNCTION

A. DERIVATION OF THE OPTIMAL INTERPOLATION COST FUNCTION

Introduction:

The formalism for the optimal interpolation technique (Lorenc 1985) is based on the Bayes theorem which relates conditional and marginal probabilities of two events. In this section, the derivation of the cost function for the optimal interpolation technique is presented. This derivation was originally presented in Lorenc (1985), but is reproduced here for completeness. We start with a prior probability of an event A, and add the information that event B has occurred, to give us a posterior probability of A given B.

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \tag{59}$$

Where *P(A)* = probability of A occurring

P(A/B) = conditional probability of A given B occurred

P(B/A) = conditional probability of B given A occurred

P(B) = probability of B occurring

For continuous variables we use probability distribution functions,

$$p(x)dx = P(x \le x^t \le x + dx)$$

where x^t is the true value. Bayes theorem takes the following form:

$$p(x/y^o) = \frac{p(y^o/x) * p(x)}{p(y^o)} \tag{60}$$

We start by thinking of x as a model state which accumulates our knowledge. The prior distribution is $p(x)$; our knowledge from previous observations. $P(x/y^o)$ is the posterior distribution, after adding the information from the previous observations y^o. $P(y^o/x)$ is the probability density of getting the observation y^o, given our previous knowledge. We further assume, our prior knowledge is that x is near x^b, and that the variance of its deviation from x^b is V^b. The usual way of modeling such a distribution is as a Gaussian:

$$p(x) = N(x / x^b, V^b) = (2\pi V^b)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(x - x^b)^2}{V^b} \right) \tag{61}$$

If our observation y^o, directly measures the variable x, with observational error V^o, then the probability of the observed value y^o, given the true value is x, can also be modeled by a Gaussian pdf:

$$p(y^o / x) = N(y^o / x, V^o) = (2\pi V^o)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(y^o - x)^2}{V^o} \right) \tag{62}$$

Using (2) and (3) $P(y^o)$ is obtained by integrating over all x:

$$\begin{aligned} p(y^o) &= \int p(y^o / x) \times p(x) \times dx \\ &= N(y^o / x^b, V^o + V^b) = (2\pi (V^o + V^b))^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(y^o - x^b)^2}{V^o + V^b} \right) \end{aligned} \tag{63}$$

Substituting (2), (3) and (4) into the Bayes Theorem in equation (1) gives:

$$p(x / y^o) = N(x / x^a, V^a) = (2\pi V^a)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(x - x^a)^2}{V^a} \right) \tag{64}$$

where,

$$x^a = V^a\left(\frac{y^o}{V^o}+\frac{x^b}{V^b}\right) \quad V^a = \left(\frac{V^o V^b}{V^o + V^b}\right)$$

Taking logarithms on both sides of (5), the Gaussians become quadratics, which are summed to give another quadratic:

$$-\ln\left[p(x/y^o)\right] = -\ln\left[p(y^o/x)\right] - \ln\left[p(x)\right] + \text{constant}$$

$$-\ln\left[p(x/y^o)\right] = \frac{1}{2}\frac{(y^o - x)^2}{V^o} + \frac{1}{2}\frac{(x^b - x)^2}{V^b} + \text{constant} \qquad (65)$$

$$= \frac{1}{2}\frac{(x^a - x)^2}{V^a} + \text{constant}$$

Therefore the posterior distribution depends on the observations and the corresponding error variances.

Bayesian formalism applied to a model:

The Bayesian analysis can be applied to assimilate observations into a model using joint pdfs and multi-dimensional Gaussian pdfs. We start by considering a model with two grid points:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Prior estimates at the grid points are given by:

$$x^b = \begin{pmatrix} x_1^b \\ x_2^b \end{pmatrix}$$

Assuming we have a prior estimate x^b with error variance V^b, the Gaussian pdf is given by:

$$p(x_1) = N(x_1 / x_1^b, V^b) = (2\pi V^b)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(x_1 - x_1^b)^2}{V^b} \right) \tag{66}$$

Since x^b is produced by a single process, usually a forecast, errors in x_1^b and x_2^b are correlated. Therefore,

$$\left\langle \left(x_1^b - x_1\right) \times \left(x_2 - x_2\right) \right\rangle = \mu V^b$$

where μ is the correlation coefficient. The corresponding joint pdf for the background forecast is modeled by a multi-dimensional Gaussian:

$$\begin{aligned} p(x_1 \cap x_2) = p(x) &= N(x / x^b, B) \\ &= \left((2\pi)^2 |B|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (x - x^b)^T B^{-1} (x - x^b) \right) \end{aligned} \tag{67}$$

where B is the covariance matrix:

$$B = V^b \begin{pmatrix} 1 & \mu \\ \mu & 1 \end{pmatrix} \tag{68}$$

We now consider the new information provided by the observation. A perfect instrument would measure a "true" value y^t. But real instruments are not perfect and have instrumental errors associated with their measurements which can be modeled by a Gaussian. The corresponding multi-dimensional Gaussian for instrumental errors is given by:

$$p(y^o / y^t) = N(y^o / y^t, E) = \left((2\pi)^2 |E|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (y^o - y^t)^T E^{-1} (y^o - y^t) \right) \tag{69}$$

where, E is the instrumental error matrix.

The above Gaussian is a function of y^t, rather than x, and therefore cannot be used in the Bayes equation. In order to do so, we need to specify the probability of a perfect instrument measuring y^t, given that the true model state is x. Since y^t is not known exactly, we have to interpolate from x to y which gives rise to an error, known as the error of representativeness. The multi-dimensional Gaussian is given by:

$$p(y^t / H(x)) = N(y^t / H(x), F)$$

$$= \left((2\pi)^2 |F|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^o - H(x))^T F^{-1}(y^o - H(x))\right) \quad (70)$$

If we assume that the instrumental error and representativeness errors are independent, we can convolve them to get a combined observational error:

$$p(y^o / x) = \int p(y^o / y^t) \times p(y^t / H(x)) dy^t$$

$$= N(y^o / H(x), E + F)$$

$$= \left((2\pi)^2 |E + F|\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^o - H(x))^T (E + F)^{-1}(y^o - H(x))\right) \quad (71)$$

Substituting the above multi-dimensional Gaussian pdfs into the Bayesian equation and taking logarithms on both sides gives:

$$-\ln\left[p(x / y^o)\right] = \frac{1}{2}(x^b - x)^T B^{-1}(x^b - x) + \frac{1}{2}(y^o - H(x))^T (E + F)^{-1}(y^o - H(x)) + \text{constant} \quad (72)$$

The equation above can be written as a cost function of the form:

$$J(x) = \frac{1}{2}(x^b - x)^T B^{-1}(x^b - x) + \frac{1}{2}(y^o - H(x))^T (E + F)^{-1}(y^o - H(x)) \quad (73)$$

So the most probable $\boldsymbol{x}$ minimizes the cost function to give us the analysis most consistent with the observations. The above equation is expressed in final form as:

$$J(\boldsymbol{x}) = \frac{1}{2}\left(\boldsymbol{x}^{\mathrm{b}} - \boldsymbol{x}\right)^{\mathrm{T}} \boldsymbol{B}^{-1}\left(\boldsymbol{x}^{\mathrm{b}} - \boldsymbol{x}\right) + \frac{1}{2}\left(\boldsymbol{y}^{\mathrm{o}} - \mathrm{H}(\boldsymbol{x})\right)^{\mathrm{T}} \boldsymbol{R}^{-1}\left(\boldsymbol{y}^{\mathrm{o}} - \mathrm{H}(\boldsymbol{x})\right) \quad (74)$$

where **B** is the Background error covariance matrix

R is the observation error covariance matrix