

Validation of the 2D-VAR lidar retrieval algorithm for non-homogeneous wind fields using FINO1 and SCADA data

Maria Krutova¹, Mostafa Bakhoday-Paskyabi¹, and Joachim Reuder¹

¹Universitetet i Bergen

June 30, 2023

Abstract

A scanning lidar measures a projection of the actual wind speed to the line-of-sight. Reconstruction of the original field is possible via various retrieval algorithms, especially if several lidars are deployed in the area. However, this may not be the case due to financial constraints. Additionally, conventional algorithms like Volume Velocity Processing (VVP) do not perform well for non-homogeneous flow fields, e.g., smooth wind turbine wakes in a retrieved field. A 2D-VAR algorithm using VVP as an intermediate step was suggested allowing the retrieval for consecutive scans at a low elevation angle. We validate this algorithm for scans containing prominent wakes under elevation angles within 5° using FINO1 mast and supervisory control and data acquisition (SCADA) data. We show that prominent wakes scanned at the hub height introduce substantial disturbances in the VVP retrieval output and affect the retrieved field. The unwanted disturbance may be overcome by masking the wakes. In point measurements, the wind speed magnitude is retrieved with an acceptable accuracy, but the wind direction appears to be sensitive to the weights and initial guess chosen. According to our sensitivity analysis, the weight assigned to the radial velocity residuals affects the outcome most, while other weights act as corrections.

RESEARCH ARTICLE

Validation of the 2D-VAR lidar retrieval algorithm for non-homogeneous wind fields using FINO1 and SCADA data

Maria Krutova* | Mostafa Bakhoday-Paskyabi | Joachim Reuder

¹Geophysical Institute and Bergen Offshore Wind Center (BOW), University of Bergen, Allégaten 77, Bergen 5007, Vestlandet, Norway

Correspondence

*Maria Krutova, Geophysical Institute and Bergen Offshore Wind Center, University of Bergen, Allégaten 77, Bergen 5007, Vestlandet, Norway Email: Maria.Krutova@uib.no

Abstract

A scanning lidar measures a projection of the actual wind speed to the line-of-sight. Reconstruction of the original field is possible via various retrieval algorithms, especially if several lidars are deployed in the area. However, this may not be the case due to financial constraints. Additionally, conventional algorithms like Volume Velocity Processing (VVP) do not perform well for non-homogeneous flow fields, e.g., smooth wind turbine wakes in a retrieved field. A 2D-VAR algorithm using VVP as an intermediate step was suggested allowing the retrieval for consecutive scans at a low elevation angle. We validate this algorithm for scans containing prominent wakes under elevation angles within 5° using FINO1 mast and supervisory control and data acquisition (SCADA) data.

We show that prominent wakes scanned at the hub height introduce substantial disturbances in the VVP retrieval output and affect the retrieved field. The unwanted disturbance may be overcome by masking the wakes. In point measurements, the wind speed magnitude is retrieved with an acceptable accuracy, but the wind direction appears to be sensitive to the weights and initial guess chosen. According to our sensitivity analysis, the weight assigned to the radial velocity residuals affects the outcome most, while other weights act as corrections.

KEYWORDS:

lidar retrieval; wind turbine wake; wind farm

1 | INTRODUCTION

Wind lidar systems, with their capability to measure the wind field remotely, have become a common measurement technology in atmospheric boundary layer research^{1,2,3}. In particular, the wind energy meteorology has utilized this measurement technology extensively during the last decade⁴. The potential of lidar application is already shown for wake dynamic studying⁵, turbulence^{6,7} and power curve measurements^{8,9} or data assimilation for a WRF model¹⁰. Processing the data in real-time opens a possibility of using lidars in short-term forecasting, e.g., for pitch¹¹ or yaw¹² control. 2D velocity fields obtained with a scanning lidar are particularly interesting for wake identification and characterization^{13,14} or validation of large-eddy simulations (LES)¹⁵.

Unlike sonic or cup anemometers, scanning lidars do not measure the wind speed directly but only its projection to the line-of-sight – the radial velocity. Hence, characterizing wakes with a scanning lidar without additional processing is effective only if the wakes are aligned with the lidar's line-of-sight. Retrieving the original wind speed and direction may be complicated when only one lidar is deployed at the site. Certain widely used algorithms require a specific scanning procedure not compatible with a 2D plane scanning. E.g., Velocity Azimuth Display (VAD) operates on conical 360° scans¹⁶. Its derivative, Doppler Beam Swinging (DBS), fetches the radial velocity from beams shoot in four directions at 30° angle from vertical. Another algorithm, Volume Velocity Processing (VVP), may be used with a scanning lidar for retrieval¹⁷ but assumes a homogeneous flow

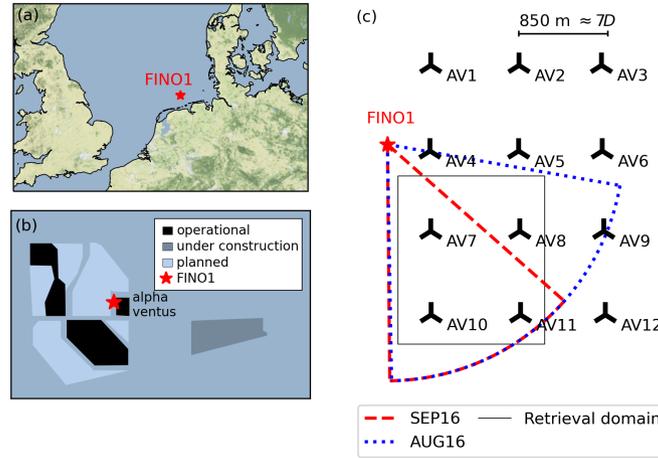


Figure 1 Map indicating the location of the FINO1 platform in the North Sea (a), the surrounding wind farms during the measurement period in September 2016 (b), and the layout of the Alpha Ventus research wind farm with the scanning angles for the PPI scans in this study (c). The retrieval is performed within the retrieval domain; here – pictured in a size used for SEP16 scans.

and does not adequately resolve heterogeneous structures, such as wakes. Complex algorithms, e.g., 4D-VAR and its derivatives, perform better on heterogeneous flows but may be time-consuming to allow real-time processing¹⁸. 4D-VAR methods are mostly developed for weather prediction models^{19,20}, their capability to estimate thermodynamic quantities may be excessive for a smaller scale wind energy research. An alternative, the 2D-VAR algorithm, was suggested to preserve advantages of the 4D-VAR method and overcome complications caused by heterogeneous wake fields²¹. The algorithm performs a retrieval from a set of consecutive lidar scans taken at a low elevation angle provided that the time interval between two scans is small enough ($\Delta t \sim 60$ s) to assume weak changes in the flow.

The 2D-VAR algorithm was validated on a one-day dataset from August 31, 2016. Due to the limited period, the validation covered only a narrow range of wind speeds (2–6.5 m/s) and directions (150–265°). Additionally, the regarded lidar scans were taken at a nearly horizontal plane and captured the flow at the bottom tip of the wind turbines. Consequently, the results could be validated only with FINO1 data. We perform a new validation on the extensive dataset spanning the whole month of September 2016. The scans from that period allow studying the algorithm's performance for a higher elevation angle. The lidar, therefore, scans one turbine at the hub height allowing a comparison to the supervisory control and data acquisition (SCADA) data for this turbine, which was not possible in the original study.

2 | DATA

2.1 | Description of the datasets

The meteorological and lidar data were obtained during the Offshore Boundary-Layer Experiment at FINO1 (OBLEX-F1) campaign in 2015–2016. The FINO1 platform is located in the North Sea at 54° 00' 53.5''N, 6° 35' 15.5''E, 45 km north of the German island of Borkum. The Alpha Ventus wind farm is located east of the FINO1 platform and consists of 12 wind turbines arranged in a rectangular pattern (Fig. 1). The wind turbines AV1–AV6 are of the type Repower 5M with a hub height of 92 m and a rotor diameter of 126 m; AV7–AV12 are of the type AREVA M5000 with a hub height of 91.5 m and a rotor diameter D of 116 m. The construction of the Borkum Riffgrund II wind farm south of the Alpha Ventus wind farm was completed in 2018 and did not affect the flow during the measurement period.

The scanning Doppler wind lidar (Leosphere WindCube 100S) was installed at the FINO1 platform at 23.5 m above sea level and was oriented towards the Alpha Ventus wind farm. The lidar operates in PPI mode. The 2D-VAR algorithm was validated on the data from a single day of August 31, 2016. On that day, the lidar performed scans under the elevation angle of 0.5°. The elevation angle remained constant for 20 hours, which resulted in nearly 1500 scans. The lidar scanned a sector between 100° and 180° azimuth angles, so that four wind turbines (AV7, AV8, AV10, AV11) and their bottom tip wakes were present in the scanned sector. We consider this one-day dataset for a comparison to the original results^{21,22} and refer to it as AUG16.

During September 2016, the lidar performed azimuth scans between 131.5° and 180° in the following elevation pattern: during the first 20 minutes of each hour, the lidar scanned at a constant elevation angle of 4.6° so that AV7 is scanned near the hub height; for the remaining 40 minutes, the lidar performed alternating scans at three elevation angles of 0.5°, 4.6°, and 9.0°. The interval between two scans is approximately 51 s.

The 2D-VAR retrieval algorithm requires scans taken at the same elevation angle with a sufficiently high temporal resolution that allows assuming weak changes in flow. Hence, we perform a lidar retrieval only for the scans taken during the first 20 minutes of each hour when the elevation angle is fixed at 4.6° . Those selected scans cover about 34% duration of September 2016. Due to an inclined scanning plane and smaller sector scanned, only the wakes from wind turbines AV7 and AV10 are always captured well in the September 2016 scans. The wakes from other wind turbines appear only for the specific wind directions. We use this one-month dataset to perform an extensive validation of the 2D-VAR algorithm and refer to it as SEP16.

Depending on weather conditions, the lidar's beam range varies and may reach up to 3 km. The far range experiences data losses frequently and thus poses little interest for the retrieval. When processing lidar data, we limit the beam range to 2.25 km for the sake of the dataset uniformity and exclusion of the area with a large percentage of missing data points.

The 10-minute averaged time series from the FINO1 mast are used as an initial guess for the retrieval algorithm and as the reference data for validation. The wind speed is measured with cup anemometers installed at 33, 40, 50, 60, 70, 80, 90, and 100 m above sea level; the mast shadow affects wind speed measurements for the wind direction of $290\text{--}330^\circ$. The wind direction is measured with vanes installed at 33, 40, 50, 60, 70, 80, and 90 m above sea level; the mast shadow affects measurements for the wind direction of $110\text{--}150^\circ$ (Fig. 2a).

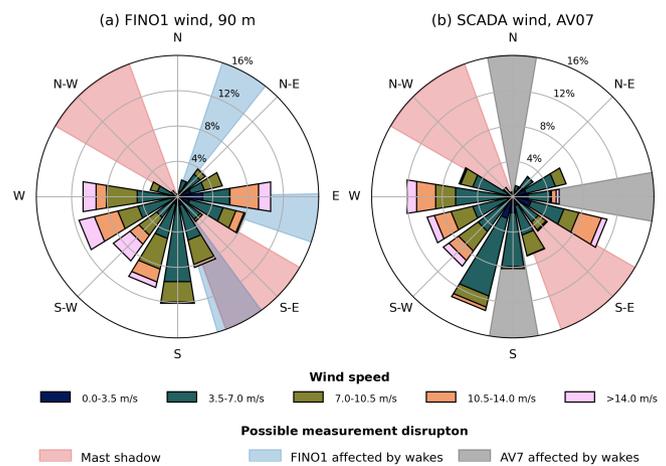


Figure 2 Wind roses for 10-minute averages of FINO1 (90–100 m measurements) and SCADA of the AV7 turbine. The wind roses consider only the periods for which both FINO1 and SCADA data are available

Since the FINO1 mast is located at a distance from the Alpha Ventus wind farm and local measurements do not represent the actual wind flow near the wind turbines, we utilize SCADA data to verify the retrieved flow near the AV7 turbine. The SCADA data consist of 1 Hz wind speed and direction time series.

FINO1 and SCADA data may differ when AV7 is affected by the wake while the FINO1 mast remains in the free flow or vice versa. The wind roses for FINO1 and SCADA data show a discrepancy in the wind speed distribution for southerly directions ($160\text{--}200^\circ$): the wind turbine AV7 is affected by the wake from AV10. Another strong difference is the easterly direction: AV7 is hit directly by the wake from AV8. Being shifted slightly to the North, the FINO1 mast is less affected by the wake from AV4. Hence, the AV7 inflow wind speeds are generally lower than registered at FINO1 (Fig. 2b). The adjacent sector of $140\text{--}160^\circ$ represents cases when the FINO1 mast is affected by the far wake from AV7, reducing the wind speed. The discrepancy between FINO1 and SCADA data there is noticeable but less prominent than for southerly and easterly winds.

The FINO1 mast data are used to define an initial guess for the retrieval algorithm. Since the FINO1 time series is nearly continuous and has little missing data in the regarded period (Table 1), we fill the gaps with a linear interpolation to obtain reference wind speed and direction values for all lidar scans. Due to larger gaps in the SCADA data and its lower importance for the 2D-VAR algorithm, we do not perform a similar interpolation for the SCADA time series. Hence, the cross-comparison for FINO1, SCADA and retrieval results (e.g., wind roses) does not consider time stamps where SCADA data are missing.

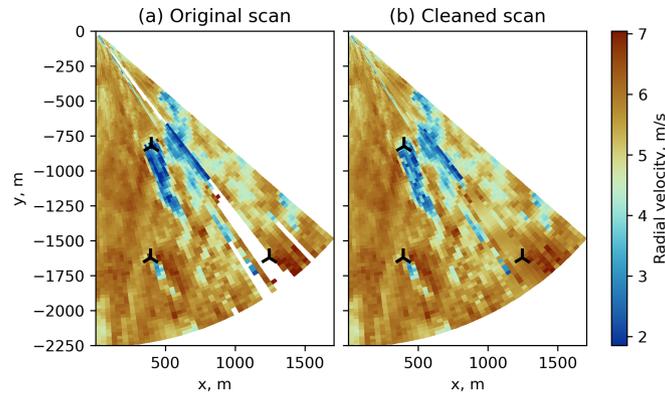
The lidar and SCADA data are recorded roughly in 1-minute intervals as opposed to 10-minute averages from FINO1. During the retrieval phase, we consider 10-minute averages from the FINO1 mast for all lidar scans in 5-minute intervals before and after the respective time stamp in the

Table 1 Data availability for the SEP16 dataset. Only scans separated by $\Delta t < 60$ s are considered.

Dataset	Subset	Missing data, %
FINO1 mast	wind speed	0.03
	wind direction	0.08
SCADA	wind speed	1.07
	wind direction	4.05
Lidar	whole scan	1.68
	retrieval domain	0.66

mast data, to comply with the higher time resolution of the lidar data. When comparing the results, we average the retrieval and SCADA series over the corresponding 10-minute periods.

2.2 | Pre-processing of the lidar scans

**Figure 3** Pre-processing of the lidar scan taken at 2016-09-05 00:08:33. a) Original scan, b) Spike removal and gap filling

The performance of the retrieval algorithm strongly depends on the lidar scan quality. The scans may contain spikes near the wind turbine position because of the lidar beam hitting the rotating blades. We identify spikes by fitting a normal, log-normal, or Weibull distributions to the radial velocity distribution within a scan. The points falling outside the best fit's confidence interval of 99.9% are regarded as spikes and removed from the scan. The gaps, both the ones in the original data set and the ones added by the spike removal, are filled with the Gaussian kernel inpainting algorithm²³ (Fig. 3b).

Occasionally, the lidar may not complete the scan correctly, resulting in missing data that cannot be restored accurately by the gap filling. A high percentage of missing data renders the scan unsuitable for the retrieval. Therefore, we only perform the lidar retrieval if more than 50% of the scanned sector is available within the retrieval domain defined as a rectangle (Fig. 1c). The discarded scans constitute less than 1% of the SEP16 dataset and occur only on certain days (Table 1). The threshold of 50% data losses is enough to ensure that the flow is resolved at least near AV7. The data around AV10 may still be missing in scans with medium data losses (20–50%). Occurrences of medium data losses are scattered across the dataset with a large cluster localized on September 19–21, characterized by low wind speed of 0–5 m/s and wind direction of 0–120° (Fig. 4).

The 2D-VAR retrieval algorithm requires an estimation of various derivatives. Due to a small azimuth step of $1^\circ \approx 0.174$ rad and the discrete nature of a scanned flow field, the derivative $\partial V_r / \partial \theta$ may return values of a higher order than the derivatives taken along the Cartesian grid or the time derivative. The derivative spikes are most likely to occur at the boundary between wakes and free-flow. These effects are mitigated by applying a Gaussian filter to the V_r field before taking a derivative.

After the cleaning, the lidar data coordinates are converted from a polar (r, θ) to a Cartesian grid (x, y) . The coordinate center is located at FINO1, with the x -axis being positive in the East direction and the y -axis being positive in the North direction. The retrieval is performed for a rectangular domain, interpolated to an equidistant Cartesian grid. The selected retrieval domain contains enough space around AV7 to resolve

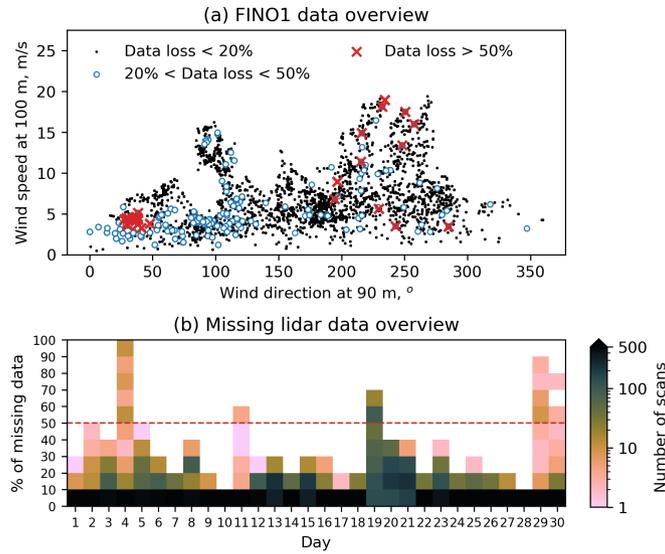


Figure 4 Overview of the missing scan data in SEP16. a) Occurrences of missing data with respect to FINO1 wind speed and direction measurements, b) number of lidar scans with a certain percentage of missing data by day.

the near wake and provide the inflow wind speed. Following the original study²¹, we perform the retrieval for AUG16 in a rectangular domain $1500 \times 1200 \text{ m}^2$ with a bottom left corner located at $x = 100 \text{ m}$, $y = -1700 \text{ m}$ and a Cartesian grid spacing of $\Delta x = \Delta y = 30 \text{ m}$. We extend the retrieval domain for SEP16 to capture an additional area around AV10 in case of northerly winds. SEP16, therefore, uses the retrieval domain of $1600 \times 1400 \text{ m}^2$ with a bottom left corner located at $x = 100 \text{ m}$, $y = -1700 \text{ m}$. The Cartesian grid spacing is increased to $\Delta x = \Delta y = 35 \text{ m}$.

3 | METHODOLOGY

The 2D-VAR algorithm²¹ consists of two steps. In the first step, the background flow is estimated with a simplified VVP retrieval algorithm. The directional wind speed components from the retrieved background flow (u_b, v_b) are then used in the second step, among with other input data, to optimize the cost function.

3.1 | VVP retrieval to estimate the background flow

The lidar measures the radial velocity V_r , i.e., the projection of the 3D wind vector on the measurement line-of-sight of the lidar:

$$V_r = u \sin \theta \cos \varphi + v \cos \theta \cos \varphi + w \sin \varphi \quad (1)$$

where u , v , and w are the directional components of U ; θ is the azimuth angle, and φ is the elevation angle of the lidar beam. The lidar scans AV7 near the hub height at an elevation angle of $\varphi = 4.6^\circ = 0.08 \text{ rad}$, which is small enough to approximate $\cos \varphi$ as 1 and $\sin \varphi$ as φ . Considering that the vertical component w is generally small, the last term of Eq. (1) can be neglected: $w \sin \varphi \approx w \varphi \approx 0$. Then, Eq. (1) is reduced to

$$V_r = u \sin \theta + v \cos \theta \quad (2)$$

The difference between the retrieved field (u, v) and the measured radial velocity V_r can be then presented as a function $F(u, v)$:

$$F(u, v) = u \sin \theta + v \cos \theta - V_r \quad (3)$$

Eq. (3) would have several solutions for $F(u, v) = 0$ if solved for a single point. Instead, we optimize $F(u, v)$ for a fetch area around the target point and find the solution via a least squares method assuming that (u, v) components remain the same within this area. The fetch area size is selected so that the variation in radial velocity is not confused with random fluctuations caused by turbulence²⁴. We pass an initial guess as directional components (u_0, v_0) calculated from FINO1 reference wind speed U and direction Φ . Since the wind direction is defined in a meteorological sense, it has to be converted to the mathematical direction as $1.5\pi - \Phi$ to obtain wind speed components via projection to x - and

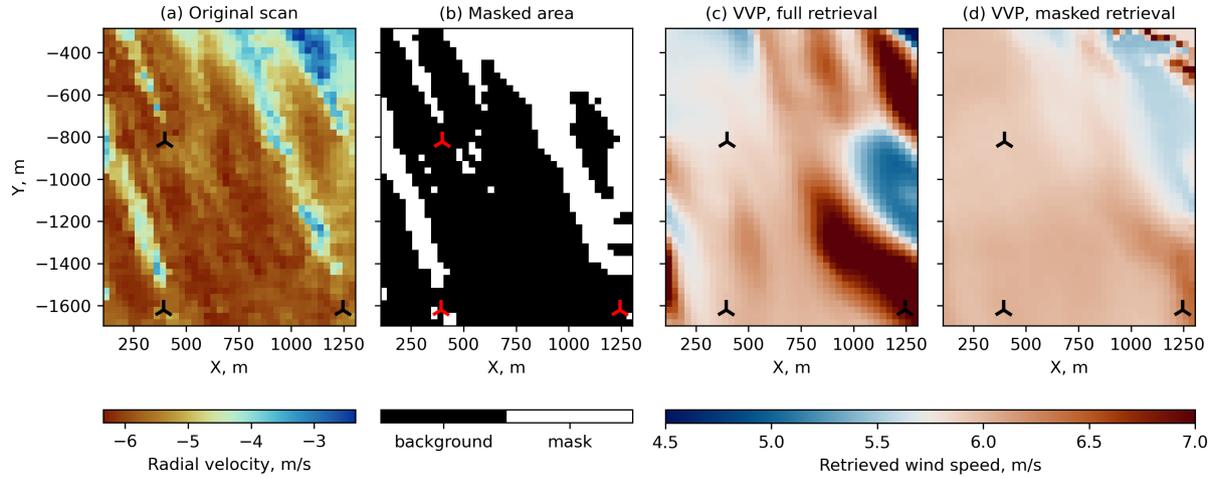


Figure 5 VVP retrieval results for an example scan from AUG16 dataset. (a) A lidar scan taken at 20160831 2:15:46 UTC+0 at the elevation angle of $\varphi = 0.5^\circ$. FINO1 wind speed 5.8 m/s, wind direction 147° . (b) Wake mask. (c) The VVP algorithm without wake masking produces structures of increased wind speed in the far range. (d) The VVP retrieved field with wake masking is more homogeneous.

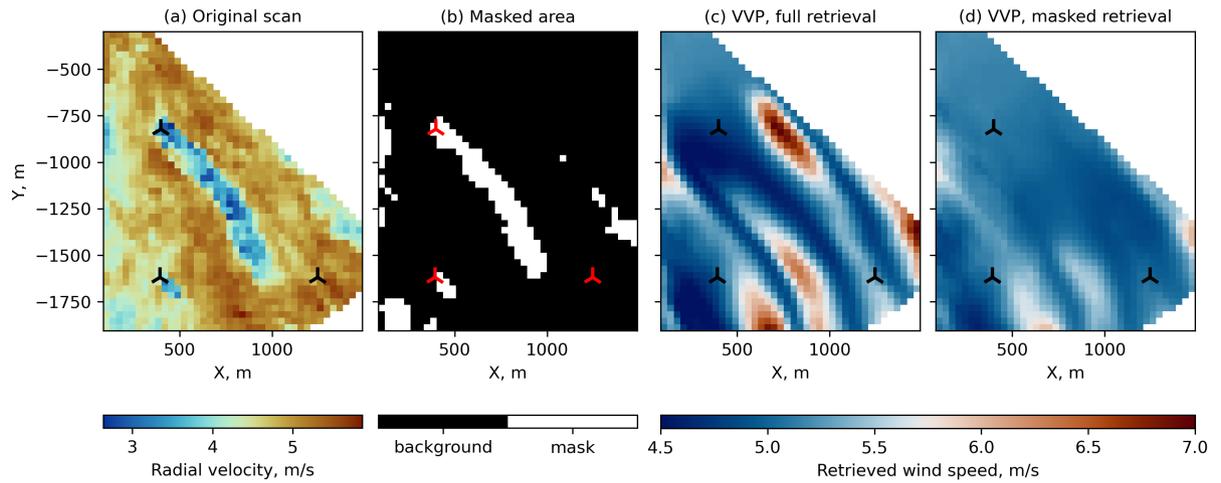


Figure 6 VVP retrieval results for an example scan from SEP16 dataset. (a) A scan taken at 20160904 21:15:22 UTC+0 at the elevation angle of $\varphi = 4.6^\circ$. FINO1 wind speed 5.3 m/s, wind direction 309° . (b) Wake mask. (c) The VVP retrieved field without wake masking is strongly heterogeneous but does not represent the actual wakes. (d) Wake masking smooths the retrieved field, particularly the high-speed structures along wakes.

y-axes.

$$u_0 = U \sin(1.5\pi - \Phi), v_0 = U \cos(1.5\pi - \Phi) \quad (4)$$

In the original study, the VVP algorithm was applied to horizontal scans at $\varphi = 0.5^\circ$, capturing the flow at the bottom tip of the wind turbines (Fig. 5a). Since the VVP algorithm assumes homogeneous wind flow but is solved locally, the retrieved field tends to overestimate the wind speed along the wakes to compensate for the wake deficit. Although the wakes captured at this level are weak and narrow, the VVP algorithm occasionally produces high-speed structures in the far range (Fig. 5c). When scanned at a higher elevation angle ($\varphi = 4.6^\circ$), the wakes occupy a larger part of the lidar scan and show a stronger wake deficit than the bottom tip wakes (Fig. 6a). Consequently, the VVP retrieval produces high-speed structures along the AV7 wake (Fig. 6c). This disturbance has to be mitigated so that it is not carried over to the next step of the 2D-VAR algorithm.

We identify and mask the wakes using an automatic thresholding algorithm that splits the radial velocity field into wake and free-flow points¹⁴. The identified wake points are not considered in the fetch area during the VVP retrieval (Fig. 5b, 6b). Since the equation $F(u, v) = 0$ is solved for

an area, running an optimization function around a masked point still fetches enough non-masked points for the solution. The masked points are then replaced with the optimization result, and the retrieved field becomes more homogeneous than the non-masked solution (Fig. 5d, 6d).

We apply masking only in the case of prominent wakes, i.e., when the reference wind speed is above the cut-in wind speed of 3.5 m/s and the lidar beam is aligned with the reference wind direction. If the angle between the lidar beam and wind direction are close to perpendicular, a so called crosswind, the radial component of the wind speed tends to zero. Despite the actual magnitude of wind speed, the crosswind wakes become less distinguishable from the free flow in a lidar scan; the radial velocity appears to be more homogeneous and poses less problem for the VVP retrieval.

Originally, the radius of the fetch area was chosen as $r = 200 \text{ m}$ ²¹ to retrieve the (u, v) field reliably when using the VVP algorithm. We increase this value to $r = 300 \text{ m}$ to ensure that the masked area is always filled in the case of wide or merging wakes.

3.2 | Cost function optimization for the detailed retrieval

The second step of the 2D-VAR algorithm optimizes a cost function J . The cost function introduces several constraints to the background flow to allow a single solution¹:

- A - the deviation between calculated and observed radial velocities defined from Eq. (5)

$$A = (u \sin \theta + v \cos \theta) - V_r \quad (5)$$

- B - the deviation from the constant wind defined via an unknown value P

$$B = (u \cos \theta - v \sin \theta) - W_{b2} \frac{\partial V_r}{\partial \theta} + P \quad (6)$$

We alter the definition of the term by adding a weight W_{b2} to allow an additional control of the derivative $\partial V_r / \partial \theta$. When $W_{b2} \neq 1$, the discrepancies with the actual derivative are then accumulated in the term P .

- C - the radial velocity advection equation. This term assumes that the radial velocity is stationary between two consecutive scans.

$$C = \frac{\partial V_r}{\partial t} + u \frac{\partial V_r}{\partial x} + v \frac{\partial V_r}{\partial y} \quad (7)$$

- D - the deviation from the background flow (u_b, v_b) , which is estimated in the previous step using the VVP algorithm.

$$D_a = u - u_b, \quad D_b = v - v_b \quad (8)$$

The weighted constraints are optimized for the whole retrieval domain defined on the Cartesian grid (x, y) .

$$J(u, v, P) = \frac{1}{2\Omega} \int (W_a A^2 + W_b B^2 + W_c C^2 + W_d D_a^2 + W_d D_b^2) d\Omega \quad (9)$$

where $W_a, W_b, W_c,$ and W_d are the weights that could be either constant or dynamic²². We utilize the suggested dynamic definition of the weights and provide a brief comparison to the constant weights when analyzing the results.

By the proposed definition, W_a is the only constant weight over the whole retrieval domain. The original study²² suggests the value of $W_a = 1$. Other weights are defined to adapt to local features of the flow. The weights, W_b and W_d , are complementary so that $W_b + W_d = 1$. They indicate whether the variability of radial velocities was accounted for.

$$W_d(x, y) = 1 - \frac{\sum (V_r^{VVP} - V_r^{obs})^2}{\sigma_{V_r}^2} \Big|_{(x, y)} \quad (10)$$

where $V_r^{VVP} = u_b \cos \theta + v_b \sin \theta$ is the radial velocity calculated from the retrieved field using Eq. (2). $\sigma_{V_r}^2$ is the variance of the measured radial velocity in the same fetch area around the point (x, y) as in the VVP retrieval, i.e., a circular area with the radius of $r = 300 \text{ m}$. The weight W_b is then $W_b = 1 - W_d$.

The weight W_c ensures that the advection equation is solved only for the points where the assumption of a stationary radial velocity would be valid. I.e., the advection distance Δr between two consecutive scans is

$$\Delta r = V_T \times \Delta t \quad (11)$$

where Δt is the time interval between the measurements at the same point in two consecutive scans, V_T is the tangential velocity

$$V_T = \frac{\partial V_r}{\partial \theta} = u_b \sin \theta + u_b \cos \theta \quad (12)$$

¹The original study aligns x -axis and respective u -component with the North direction. To comply with the meteorological notation where x -axis is aligned with the East, we had to switch u and v -components where it was relevant

Then, taking λ as 1–2 times the grid spacing, the weight W_c is defined as a binary matrix

$$W_c = \begin{cases} 0; & \text{for } \Delta R > \lambda \\ 1; & \text{for } \Delta R \leq \lambda \end{cases} \quad (13)$$

We observed an improvement of the retrieved flow structure in SEP16 if more weight is given to W_a , which we discuss further in the sensitivity analysis of the 2D-VAR algorithm. We also set $W_{b2} = 1$ and apply a Gaussian filter with the standard deviation of $\sigma = 4$ to the gap-filled radial velocity V_r prior of calculation of the derivative $\partial V_r / \partial \theta$ in Eq. (6). Time and spatial derivatives in the advection component Eq. (7) are calculated for the original field with the gaps filled to avoid over-smoothing of the input data. Hence, we define our weight set as follows

$$W_a = 5, W_b = 1 - W_d, W_{b2} = 1, W_c \text{ as in Eq. (13), } W_d \text{ as in Eq. (10)} \quad (14)$$

Since the cost function is optimized for the whole area at once and attempts to retrieve the original non-homogeneous flow, the wake masking is not applied here. An initial guess utilizes a constant field (u_0, v_0) as defined in Eq. (4) from the FINO1 mast data. The initial deviation from the constant flow is set to zeroes matrix $\mathbf{P}_0 = 0$.

3.3 | Validation

We consider two aspects when evaluating the retrieval accuracy: whether the virtual lidar field modeled from the retrieved flow correspond to the FINO1 lidar measurement and whether the local retrieved wind characteristics agree with FINO1 and SCADA time series.

The modelled radial velocity V_r^{mod} is calculated from the retrieved field (u^{ret}, v^{ret}) using Eq. (2). The divergence from the observed radial velocity V_r^{obs} is evaluated based on the mean and standard deviation of the residuals ΔV_r .

$$\Delta V_r = V_r^{obs} - V_r^{mod} = V_r^{obs} - (u^{ret} \cos \theta + v^{ret} \sin \theta) \quad (15)$$

The retrieved wind speed and direction are compared against FINO1 and SCADA data. We use different approaches depending on the lidar's elevation angle. For the nearly horizontal scans in the AUG16 dataset taken at the elevation angle of 0.5° , only the comparison to FINO1 measurements near the height of 33 m would be relevant. The wind speed components $(u_{xy}^{ret}, v_{xy}^{ret})$ are extracted from the top left corner of the retrieval domain – the probe point closest to the FINO1 mast, excluding the border points, or $x_p = 130$ m, $y_p = -330$ m.

The SEP16 dataset with scans taken at the elevation angle of 4.6° provides a different option for the comparison – the time series at AV7 hub height. The inflow wind speed and direction for the AV7 turbine are probed dynamically so that the probe point is always located at $1D$ upstream of AV7. The wind speed components are calculated as mean values in a circular area with a radius of $1D$ around the probe point.

For the comparison, the retrieval and SCADA time series are averaged to the same 10-minute periods as the FINO1 time series. The root mean square errors (RMSE) of wind speed ΔU and direction $\Delta \phi$ estimation are then calculated as follows:

$$\Delta U = \sqrt{\frac{1}{N} \sum_{i=1}^N (U_1 - U_2)^2}, \quad U_{(1,2)} = \sqrt{u_{(1,2)}^2 + v_{(1,2)}^2} \quad (16)$$

$$\Delta \phi = \sqrt{\frac{1}{N} \sum_{i=1}^N \left[\arccos \left(\frac{u_1 u_2 + v_1 v_2}{U_1 \cdot U_2} \right) \right]^2} \quad (17)$$

where the indices (1, 2) mark values from respective time series U_1 and U_2 and N is the number of valid values.

4 | RESULTS

4.1 | Comparison to the original study, August 31, 2016 (AUG16)

Since we are implementing the 2D-VAR retrieval algorithm anew, we verify how our implementation performs compared to the original study^{21,22}. The 2D-VAR retrieved time series follows the FINO1 time series regardless of the wake masking. In contrast, the VVP series deviate stronger from FINO1 for certain wind directions when the retrieval is run without wake masking (Fig. 7).

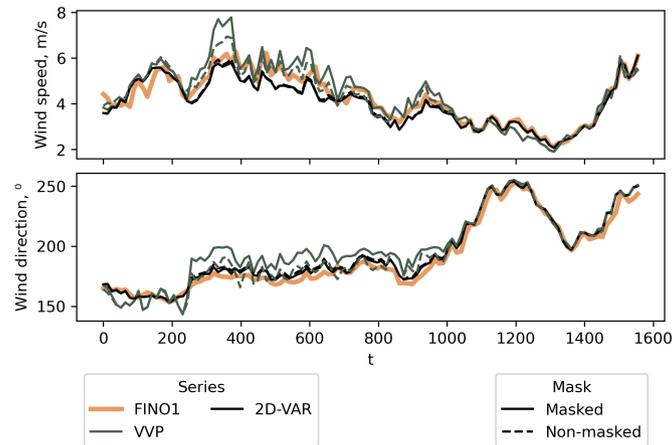
The new implementation achieves good agreement for scans taken after 15:00. Those scans are characterized by crosswind directions and wind speeds close to cut-in wind speed. I.e., the flow is nearly homogeneous. Consequently, the new implementation and VVP retrieval show little divergence from the reference data and from each other.

We observe similar errors in the wind speed estimation between our and original implementation, but a larger error in the wind direction (Table 2). The wind direction error is also larger for the VVP retrieval. Considering that the VVP algorithm returns a rather stable solution, the difference in wind direction errors may be caused by different pre-processing and solving procedures that were not detailed in the original study.

Table 2 Results reproduction for AUG16 dataset.

Study	Algorithm	Wind speed error	Wind speed correlation coefficient	Wind direction error	Wind direction correlation coefficient
Original†	2D-VAR	0.383 m/s	0.96	-1.4°	0.98
	VVP	0.29 m/s	0.98	4.3°	0.99
Reproduced	2D-VAR	0.334 m/s	0.96	6.06°	0.96
	VVP	0.375 m/s	0.96	7.61°	0.96

†Cherukuru, N. W., Calhoun, R., Krishnamurthy, R., Benny, S., Reuder, J. and Flügge, M.: 2D VAD single Doppler lidar vector retrieval and its application in offshore wind energy, *Energy Procedia*, 137, 497–504, 2017.

**Figure 7** Comparison of the retrieval results and FINO1 data for the AUG16 dataset.

4.2 | Comparison to the new dataset, September 2016 (SEP16)

SEP16 dataset provides an additional possibility for the validation provided that SCADA data are available. FINO1 vs. SCADA comparison is then taken as a reference – the retrieved series should not perform worse than it.

The retrieval time series shows good agreement with FINO1 and SCADA data, especially for the wind speed estimation (Fig. 10ab). Several outliers in wind direction cannot be explained only by being close to 0° or 360°. When highlighted by the wind speed value, the irregular outliers correspond to wind speeds below 1 m/s. The uncertainty of the instantaneous measurements can then explain the discrepancy during low wind. The directional components both in measured and retrieved time series are prone to uncertainty error. Consequently, the wind direction may be estimated incorrectly for low wind even if the wind speed magnitude is similar. Hence, we consider the retrieval results unreliable for FINO1 or SCADA wind speeds below 1 m/s and do not include them in the comparison and error calculation.

The scatter of retrieved wind direction is smaller for the 2D-VAR vs. FINO1 comparison than for 2D-VAR vs. SCADA (Fig. 10ab). In addition, the scatter in 2D-VAR vs. SCADA plot resembles FINO1 vs. SCADA comparison (Fig. 10c) despite the values being probed closer to AV7, SCADA measurement location, than FINO1. A better agreement with FINO1 implies that the cost function optimization may be tuned to the initial guess (u_0, v_0) calculated from FINO1 data. We perform a sensitivity analysis to determine, which factors affect the 2D-VAR algorithm.

4.2.1 | Sensitivity to the input data

The 2D-VAR algorithm consists of two optimization steps. Both steps require an initial guess of the horizontal wind field (u_0, v_0) . In the first step, the VVP retrieval is run locally point by point. The VVP solution via least-squares minimization is rather stable and does not noticeably depend on an initial guess. The solution can be affected slightly by altering the grid spacing or retrieval domain size. Stronger alteration can be achieved if some points are deliberately excluded from the minimization, i.e., by applying the wake masking (Fig. 11).

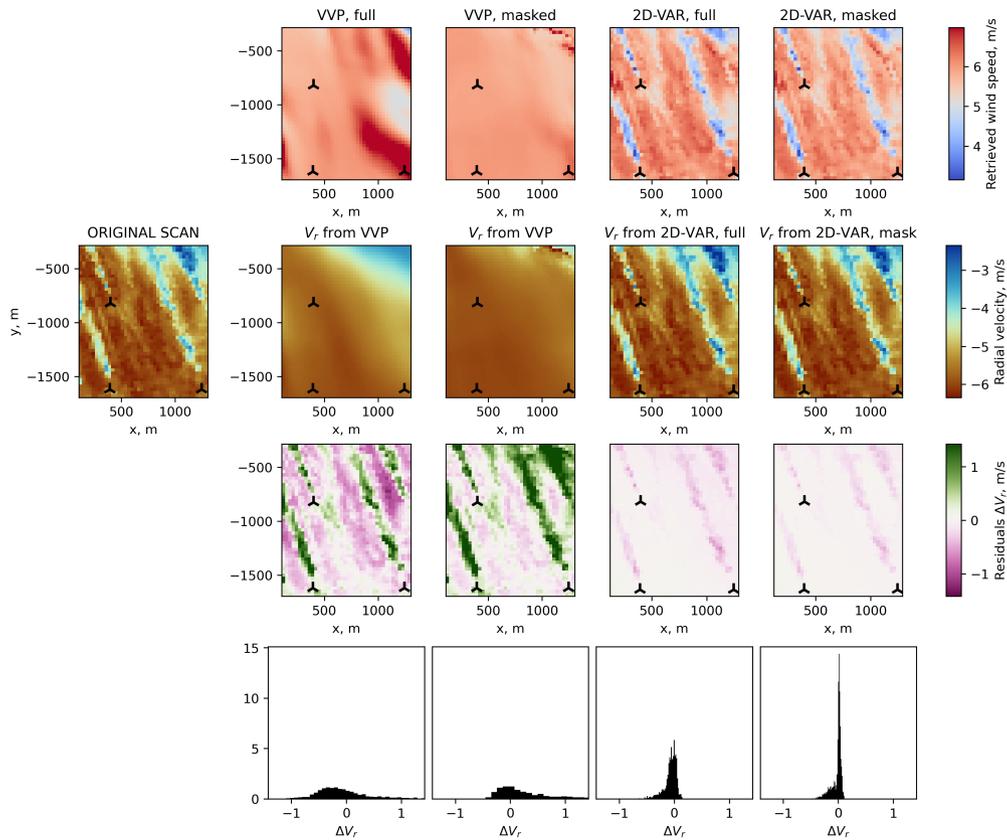


Figure 8 Example of the retrieval performed for the scan from AUG16 taken at 20160831 02:15:46 UTC+0. FINO1 wind speed 5.6 m/s, wind direction 155°.

The VVP field is then used in the cost function optimization in the term D (Eq. (8)) and the weight W_d , if the dynamic definition Eq. (10) is applied. Hence, strong non-physical disturbances in the VVP solution are carried over to the next optimization step. Figure 11 shows an example of the retrieval performed for a lidar scan with long prominent wakes present. If the VVP retrieved field is obtained without the wake masking, the 2D-VAR solution gets a non-uniform wind speed increase along the wind turbines. Consequently, radial velocity residuals have a higher standard deviation and bias compared to the solution where the wakes were masked for the VVP algorithm. The residuals of the non-masked solution tend to have lower bias and standard deviation.

For the first step of the 2D-VAR algorithm, the accuracy of wake masking becomes more important for the VVP solution than the initial guess. If the wakes cannot be identified reliably, e.g., in the crosswind conditions, it is preferable to proceed without masking. Wake masking is the most efficient when the lidar captures wide and long strong wakes, i.e., when the lidar beam and wind direction are aligned (Fig. 11). Due to the FINO1 relative position to AV7 and with the respect to the lidar orientation, such alignment occurs only in the north-western and south-eastern sectors. Although not many events are registered for those sectors, the agreement to the SCADA wind rose (Fig. 9b) is better when the wake masking is applied (Fig. 9c) as opposed to non-masked wakes (Fig. 9d). The improvement leads to a slight reduction of RMSE when the wake masking is applied compared to non-masked retrieval (Fig. 12).

The second step of the 2D-VAR algorithm relies on optimizing the cost function calculated over the whole area. Despite the 2D-VAR algorithm using the FINO1 data as an initial guess, the retrieved data's wind speed and direction distribution (Fig. 9c) are closer to the SCADA wind rose rather than FINO1. If the algorithm is run with the SCADA data as an initial guess, the agreement between the SCADA data and retrieved inflow gradually increases (Table 3, Fig. 12). On the contrary, the solution diverges from FINO1 mast data at the hub height showing wind direction errors similar to the comparison between FINO1 and SCADA.

Nevertheless, the wind field estimated with FINO1 has a smaller gap between wind direction errors for FINO1 or SCADA data comparison. A good agreement implies that the retrieved wind field near AV7 tends to the actual wind field, provided that the initial guess is still close to it.

Table 3 Validation of the retrieval time series for the SEP16 dataset. Bold font marks the smallest error of wind speed and direction.

Series	Wind speed error (RMSE)		Wind direction error (RMSE)	
	VVP	2D-VAR	VVP	2D-VAR
FINO1 mast data as initial guess				
Retrieval vs. FINO1 (60 m)	0.89 m/s (0.976)	0.80 m/s (0.981)	12.16° (0.975)	10.08° (0.980)
Retrieval vs. FINO1 (hub height)	1.04 m/s (0.967)	0.67 m/s (0.988)	14.51° (0.981)	11.68° (0.976)
Retrieval vs. SCADA	1.22 m/s (0.954)	1.08 m/s (0.976)	7.46° (0.988)	7.46° (0.988)
SCADA AV7 data as initial guess				
Retrieval vs. FINO1 (60 m)	0.88 m/s (0.975)	1.13 m/s (0.976)	12.71° (0.975)	12.23° (0.975)
Retrieval vs. FINO1 (hub height)	1.03 m/s (0.966)	1.08 m/s (0.976)	15.16° (0.976)	13.74° (0.977)
Retrieval vs. SCADA	1.31 m/s (0.949)	0.74 m/s (0.980)	8.44° (0.976)	5.85° (0.983)
Reference comparison				
FINO1 vs. SCADA	1.43 m/s (0.961)		12.69° (0.979)	

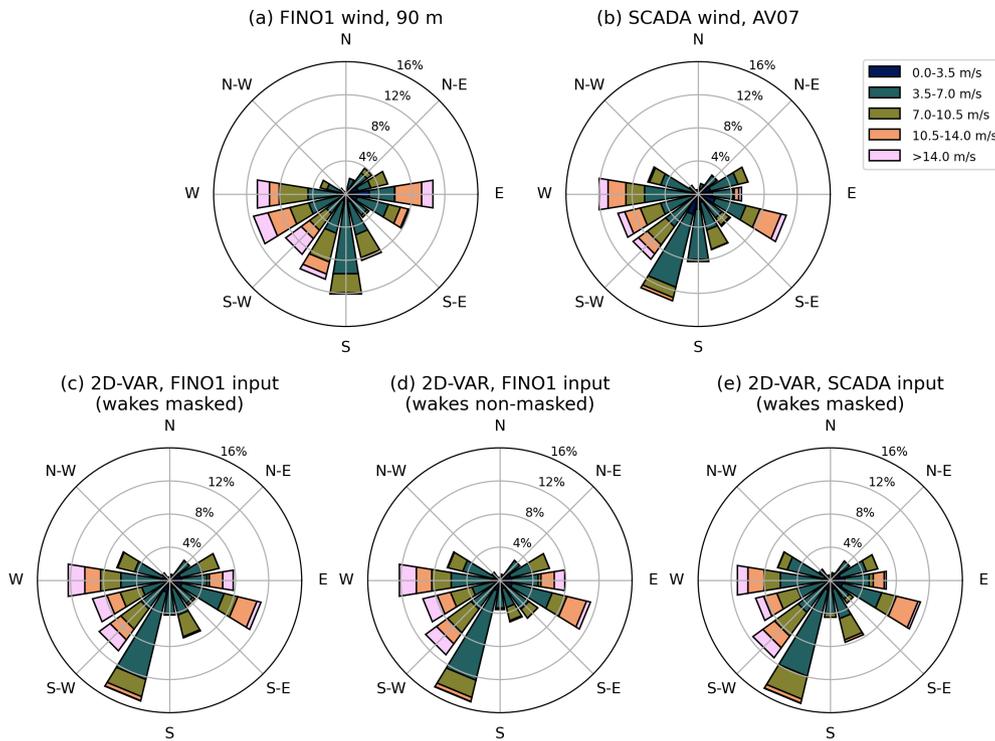


Figure 9 Wind roses for SEP16 for the inflow probe 1D upstream of AV7, 2D-VAR input sensitivity. Only scans valid for retrieval and are considered. a) FINO1 data, b) SCADA data, c) 2D-VAR retrieval, FINO1 data as initial guess, wakes are masked at VVP step, d) 2D-VAR retrieval, FINO1 data as initial guess, wakes are not masked at VVP step, e) 2D-VAR retrieval, SCADA data as initial guess, wakes are masked at VVP step

4.2.2 | Sensitivity to the weights

We alter the weights W_a , W_{b2} , W_c , and W_d to analyze their effect on the final solution. We do not regard the weight W_b separately since it is linked to W_d .

The weight W_a adjusts the cost function term A , which reduces the discrepancy between the lidar data and radial velocity calculated from the retrieved field as defined in Eq. (5). Empirically, this weight should remain at $W_a = 1$. However, we observed a non-physical wind speed increase at the wake boundaries in the retrieved flow when using this value for SEP16 scans. Increasing the weight puts more emphasis on reducing the

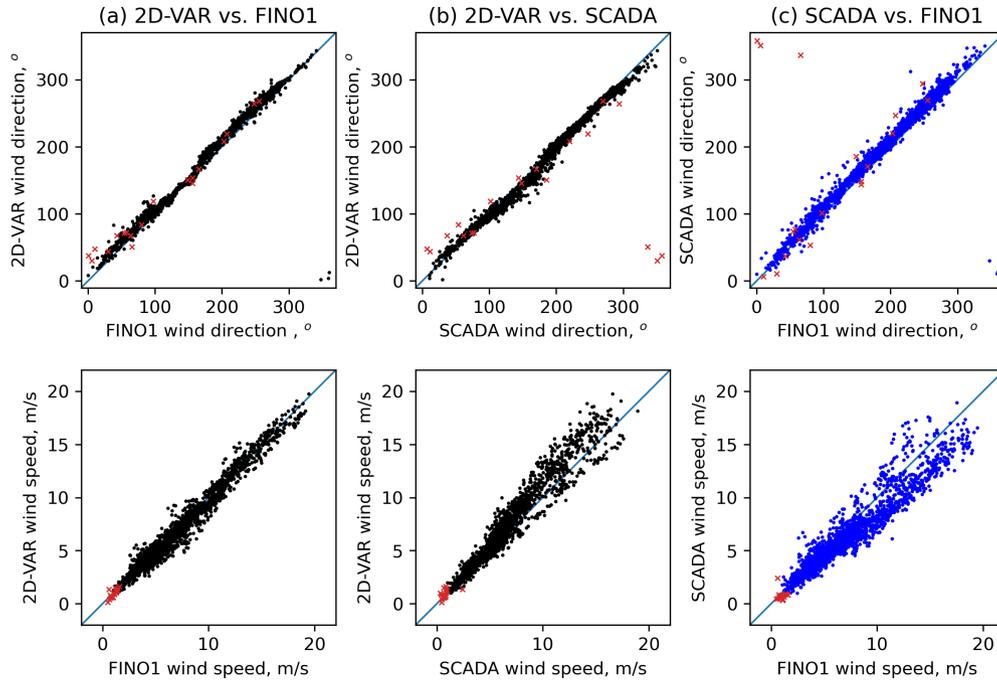


Figure 10 Wind speed and direction comparison for the 2D-VAR retrieval data at AV7 inflow. a) 2D-VAR vs. FINO1, b) 2D-VAR vs. SCADA, c) SCADA vs. FINO1

discrepancy in radial velocity residuals but leads to a bias in the residuals. The residuals for $W_a = 1$ are more symmetrical than with the increased weight and have a higher standard deviation (Fig. 13). While a further increase of W_a gradually decreases the residuals, it strongly affects wind roses for the AV7 inflow (Fig. 14ef), and they diverge from the reference FINO1 and SCADA wind roses (Fig. 14ab). Considering RMSE and correlation coefficient trends for $W_a < 10$ and $W_a \geq 10$, we suggest that using $W_a = 1$ is preferable to reduce the bias in radial velocity residuals, although the retrieved flow may occasionally produce small irregular structures. $W_a = 5$ is a good choice to remove those structures at the cost of increasing the bias in wake and free-flow residuals. We do not recommend using higher values of W_a . As shown in Fig. 13, the high weight of $W_a = 20$ may force the algorithm to stay at the initial guess field, because other components would not contribute comparably to the cost function. Small weights $W_a < 1$ should not be considered, as they lead to increased residuals.

We introduced an additional weight W_{b2} to the term B of the cost function Eq. (6). This weight may be left at $W_{b2} = 1$, provided the radial velocity field was filtered before estimating a derivative $\partial V_r / \partial \theta$. If the original radial velocity field is used, leaving the weight at $W_{b2} = 1$ results in erroneous retrieval due to the high value of derivative (Fig. 16). Decreasing the weight to $W_{b2} = 0.1$ performs nearly equally for filtered and original radial velocities but slightly increases noisiness in the retrieved field. The residuals distribution in the case of reduced weight generally remains similar to the base case: $W_{b2} = 1$ and filtered radial velocity.

Being complementary, the weights W_b and W_d behave in a similar way. We alter the weight W_d and calculate the other weight as $W_b = 1 - W_d$. Effectively, the dynamic weight W_d defined in Eq. (10) acts similarly to the wake masking while also providing local weighting for the background flow. If Eq. (10) cannot be implemented, the wake mask can be re-used in the cost function optimization. However, the wakes cannot be identified reliably in the crosswind with the current masking algorithm. W_d may be set to a constant value as an alternative. Constant weight $W_d = 0.5$ provides an equal weighting to the wake and free-flow points and slightly decreases the bias Eq. (10). Setting $W_d = 1$ may shift the bias within the wakes but does not remove it completely.

The weight W_c is defined dynamically as a binary matrix. Since W_c strongly depends on the flow speed and grid resolution, the matrix may be filled with ones in the case of the low radial velocity, e.g., weak actual flow or crosswind. Setting W_c to a constant value of one for all cases does not strongly alter the final solution. Moreover, W_c would have to be increased by several orders for the residuals to become noticeably affected.

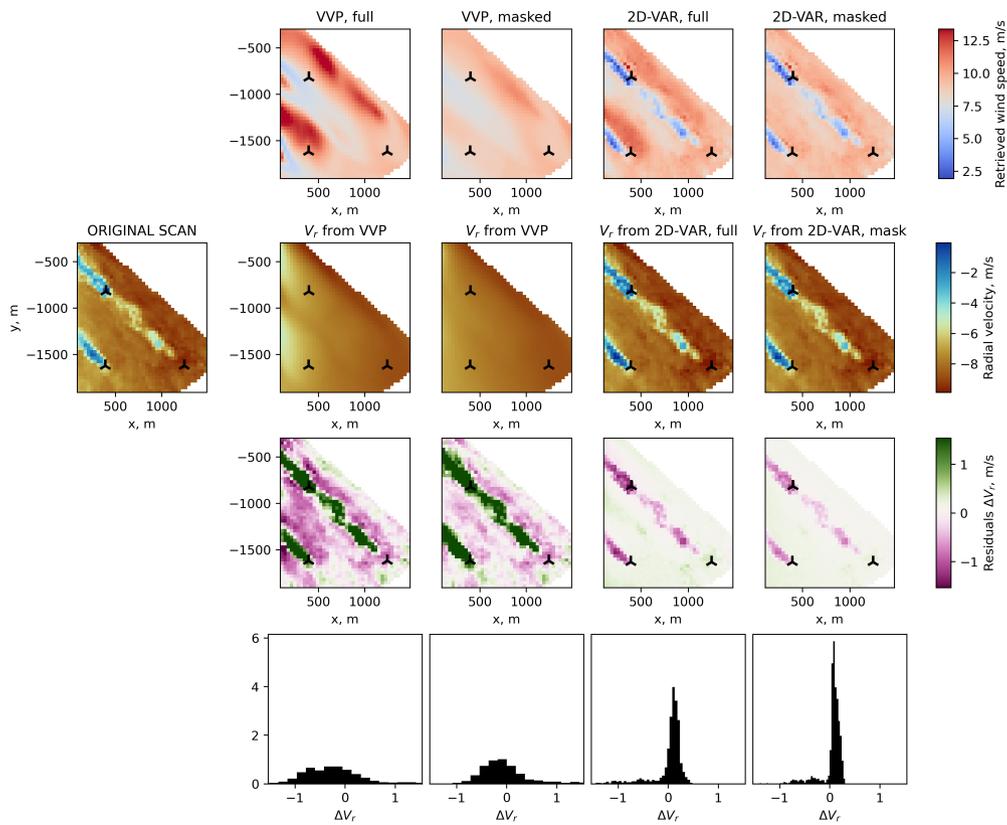


Figure 11 Example of the retrieval performed for the scan from SEP16 taken at 20160911 23:00:03 UTC+0. FINO1 wind speed 8.8 m/s, wind direction 121°.

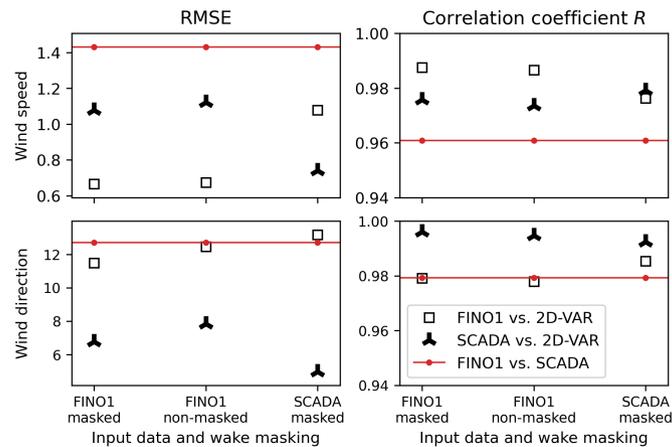


Figure 12 Wind speed and direction RMSE depending on an initial guess of the cost optimization function and wake processing for the VVP algorithm

5 | CONCLUSIONS

We independently reproduced the 2D-VAR algorithm for lidar retrieval for the August 31, 2016 dataset. While we got a similar error in the wind speed estimation, our implementation returned higher offset in the wind direction compared to the original study. We attribute the mismatch to different pre-processing procedures and optimization algorithms used and the overall sensitivity of the wind direction to the retrieval.

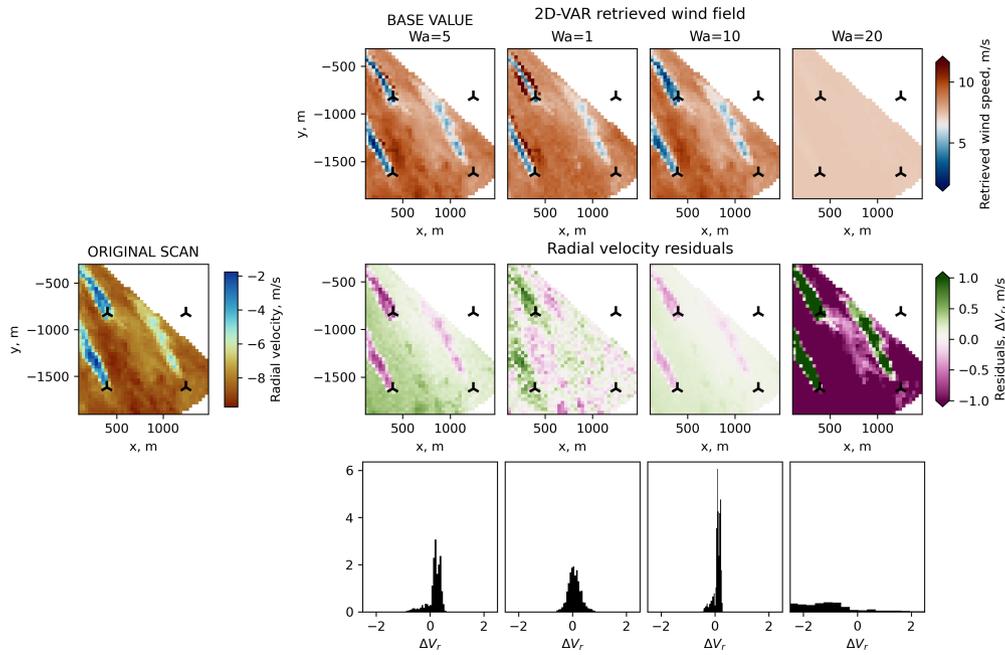


Figure 13 2D-VAR retrieved field and radial velocity residuals depending on the choice of weight W_α . Scan taken at 20160912 2:07:43 UTC+0. FINO1 wind speed 7.2 m/s, wind direction 138°.

We performed an extended validation with the September 2016 dataset and SCADA data for the same period. Due to a higher lidar elevation angle, the scans were capturing wakes near the hub height. We observed an increased heterogeneity in the first step of the 2D-VAR algorithm – an estimation of the background flow with the VVP algorithm. The effect primarily appears in the case of large and strong wakes, which are observed when the wind direction is aligned with the lidar beam. The unwanted heterogeneity was mitigated by masking the wakes with an automatic thresholding algorithm to exclude them from the VVP solution.

Besides the background flow from the VVP algorithm, the cost function optimization in the second step of 2D-VAR algorithm is found to be sensitive to weights and the initial wind field. Defining the initial wind field based on FINO1 or SCADA data tunes the resulting field to the initial values. Nevertheless, the retrieval algorithm tends to the actual flow in both cases – the agreement between wind direction retrieved near AV7 and SCADA series is always better than to FINO1 data. FINO1 data can still be used as an initial guess, when SCADA data are not available. However, if both datasets are accessible, using SCADA data becomes preferable to reconstruct the wind field near the corresponding wind turbine.

Of the weights regarded, the weight W_α is directly connected to the radial velocity residuals and, therefore, affects the residuals and the flow structure most. Other main weights – W_b , W_c , and W_d – have weaker effect on the retrieval result, although they may cause local changes in the retrieved flow. A supplementary weight W_{b2} was introduced primarily for an additional control over the $\partial V_r / \partial \theta$ derivative. Smooth derivative along azimuth θ is more important than W_{b2} in the case of a small azimuth step.

The retrieved flow behaves differently when the wakes are parallel or perpendicular. It is possible that the accuracy of 2D-VAR retrieval may be increased by adjusting the cost function depending on the wind direction relative to the scanned azimuths. Since a comparison to the 'true' flow is complicated for the lidar data, we plan to explore the 2D-VAR algorithm performance on LES generated wakes and virtual lidar in order to improve the retrieval result.

ACKNOWLEDGMENTS

The OBLEX-F1 field campaign was performed under the Norwegian Centre for Offshore Wind Energy (NORCOWE), funded by the Research Council of Norway (RCN) under project number 193821. The scanning Doppler wind lidar system (Leosphere WindCube 100S) has been made available via the National Norwegian infrastructure project OBLO (Offshore Boundary Layer Observatory) also funded by RCN under project number 227777.

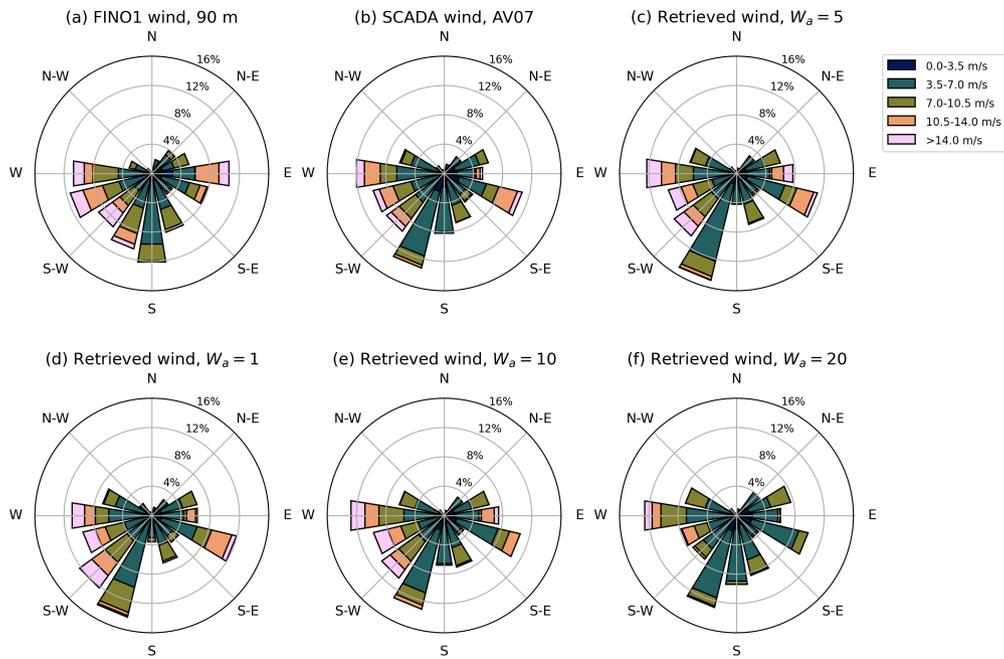


Figure 14 Wind roses for inflow probe at AV7, W_a sensitivity. Only scans valid for retrieval are considered. a) FINO1 data, b) SCADA data, c) 2D-VAR retrieval, $W_a = 1$, d) 2D-VAR retrieval, $W_a = 5$ - default parameter in this study, e) 2D-VAR retrieval, $W_a = 10$

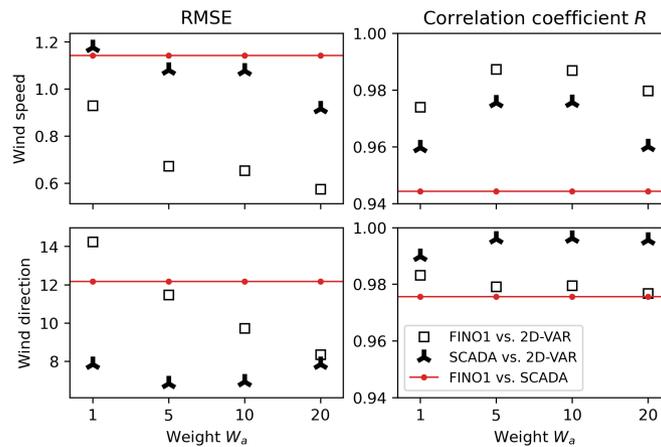


Figure 15 Wind speed and direction RMSE depending on the choice of weight W_a

Wind speed and direction data at FINO1 were made available by the RAVE (Research at Alpha Ventus) initiative, which was funded by the German Federal Ministry of Economic Affairs and Energy on the basis of a decision by the German Bundestag and coordinated by Fraunhofer IWES (see: www.rave-offshore.de).

This work was partly supported by the Research Council of Norway through "Research on smart operation control technologies for offshore wind farms (CONWIND)" project, ENERGIX, with grant number 304229.

References

1. Lothon M, Lenschow DH, Mayor SD. Coherence and scale of vertical velocity in the convective boundary layer from a Doppler lidar. *Boundary-Layer Meteorology* 2006; 121: 521-536. doi: 10.1007/S10546-006-9077-1/METRICS

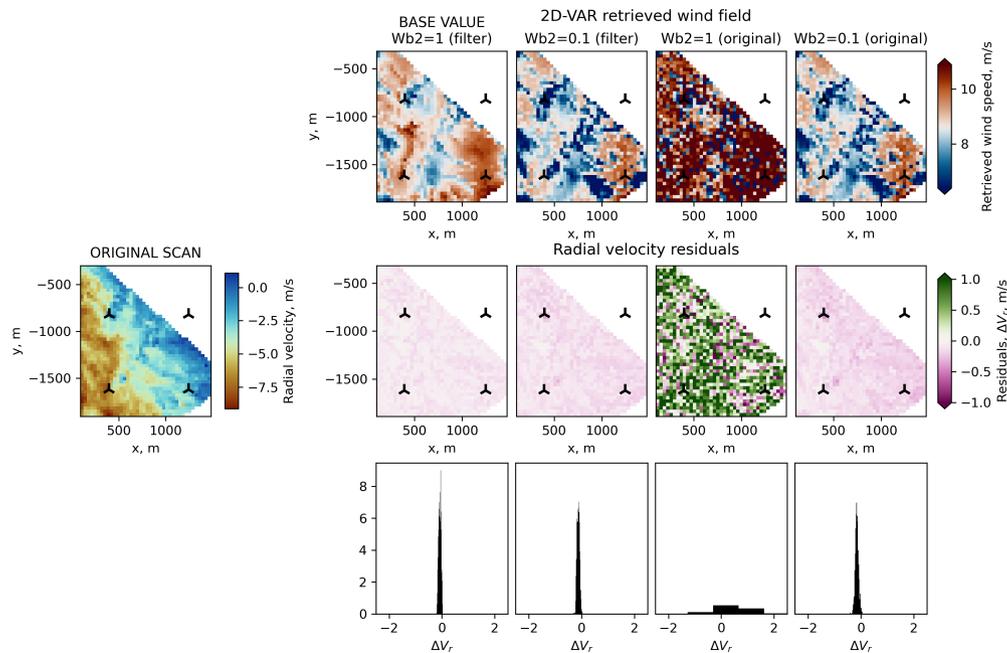


Figure 16 2D-VAR retrieved field and radial velocity residuals depending on the choice of weight W_{b2} in Eq. (6). Scan taken at 20160911 00:01:45 UTC+0. FINO1 wind speed 9.7 m/s, wind direction 201°.

2. Manninen AJ, Marke T, Tuononen M, O'Connor EJ. Atmospheric Boundary Layer Classification With Doppler Lidar. *Journal of Geophysical Research: Atmospheres* 2018; 123: 8172-8189. doi: 10.1029/2017JD028169
3. Liu Z, Barlow JF, Chan PW, et al. A Review of Progress and Applications of Pulsed Doppler Wind LiDARs. *Remote Sensing* 2019, Vol. 11, Page 2522 2019; 11: 2522. doi: 10.3390/RS11212522
4. Reuder J, Cheynet E, Clifton A, et al. Recommendation on use of wind lidars. <https://doi.org/10.5281/zenodo.4672351>; 2021
5. Trabucchi D, Trujillo JJ, Schneemann J, Bitter M, Kühn M. Application of staring lidars to study the dynamics of wind turbine wakes. *Meteorologische Zeitschrift* 2015; 24: 557-564. doi: 10.1127/METZ/2014/0610
6. Courtney M, Wagner R, Lindelöw P. Testing and comparison of lidars for profile and turbulence measurements in wind energy. *IOP Conf. Ser. Earth Environ. Sci.* 2008; 1(1): 012021. doi: 10.1088/1755-1315/1/1/012021
7. Mann J, Sathe A, Gottschall J, Courtney M. Lidar Turbulence Measurements for Wind Energy. In: . 141. Springer Science and Business Media, LLC. 2012 (pp. 263-270)
8. Wagner R, Courtney M, Gottschall J, Lindelöw P. Improvement of power curve measurement with lidar wind speed profiles. In: EWEC 2010 Proceedings online. European Wind Energy Association (EWEA); 2010.
9. Krishnamurthy R, Choukulkar A, Calhoun R, Fine J, Oliver A, Barr KS. Coherent Doppler lidar for wind farm characterization. *Wind Energy* 2013; 16: 189-206. doi: 10.1002/WE.539
10. Bakhoday-Paskyabi M, Flügge M. Predictive Capability of WRF Cycling 3DVAR: LiDAR Assimilation at FINO1. *Journal of Physics: Conference Series* 2021; 2018: 012006. doi: 10.1088/1742-6596/2018/1/012006
11. Raach S, Schlipf D, Haizmann F, Cheng PW. Three Dimensional Dynamic Model Based Wind Field Reconstruction from Lidar Data. *Journal of Physics: Conference Series* 2014; 524: 012005. doi: 10.1088/1742-6596/524/1/012005
12. Dhiman HS, Deb D, Muresan V, Balas VE. Wake Management in Wind Farms: An Adaptive Control Approach. *Energies* 2019, Vol. 12, Page 1247 2019; 12: 1247. doi: 10.3390/EN12071247

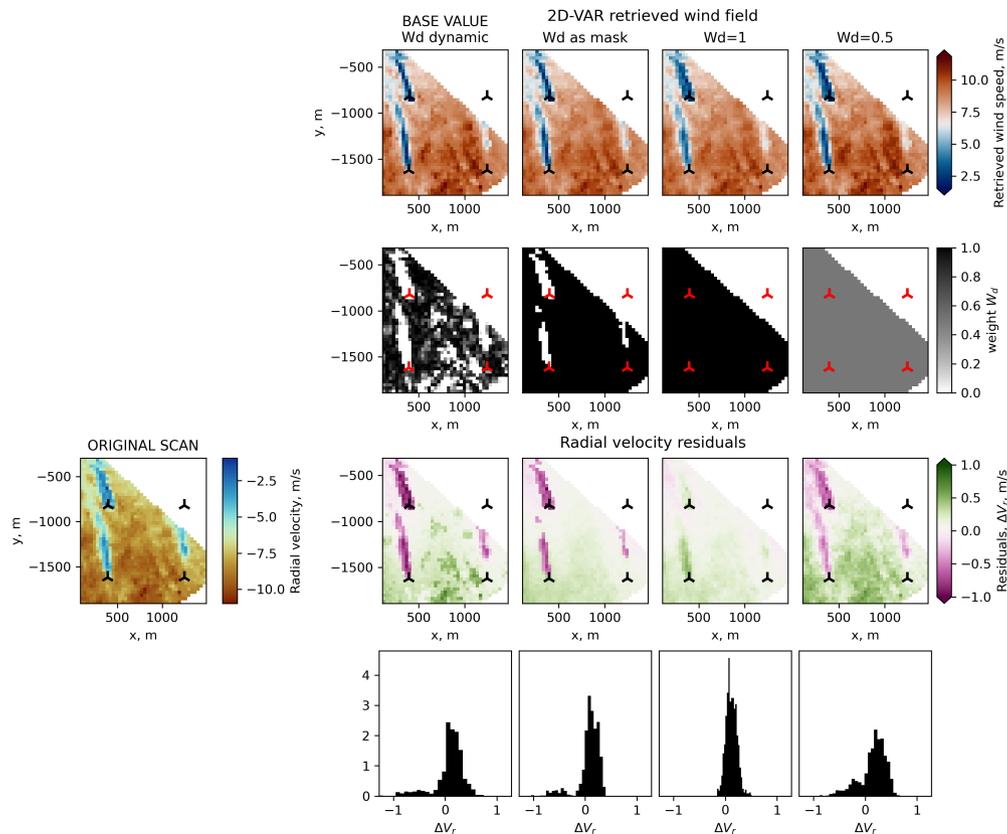


Figure 17 2D-VAR retrieved field and radial velocity residuals depending on the choice of weight W_d . Scan taken at 20160924 21:18:21 UTC+0. FINO1 wind speed 7.4 m/s, wind direction 158°.

13. Krishnamurthy R, Reuder J, Svoldal B, Fernando HJ, Jakobsen JB. Offshore Wind Turbine Wake characteristics using Scanning Doppler Lidar. *Energy Procedia* 2017; 137: 428-442. doi: 10.1016/j.egypro.2017.10.367
14. Krutova M, Bakhoday-Paskyabi M, Reuder J, Nielsen FG. Development of an automatic thresholding method for wake meandering studies and its application to the data set from scanning wind lidar. *Wind Energy Sci.* 2022; 7(2): 849-873. doi: 10.5194/wes-7-849-2022
15. Smalikho IN, Banakh VA, Pichugina YL, et al. Lidar Investigation of Atmosphere Effect on a Wind Turbine Wake. *Journal of Atmospheric and Oceanic Technology* 2013; 30: 2554-2570. doi: 10.1175/JTECH-D-12-00108.1
16. Browning KA, Wexler R. The Determination of Kinematic Properties of a Wind Field Using Doppler Radar. *J. Appl. Meteorol.* 1968; 7(1): 105-113. doi: 10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2
17. Wang H, Barthelmie RJ, Clifton A, Pryor SC. Wind Measurements from Arc Scans with Doppler Wind Lidar. *J. Atmos. Ocean. Technol.* 2015; 32(11): 2024-2040. doi: 10.1175/JTECH-D-14-00059.1
18. Janisková M. Assimilation of cloud information from space-borne radar and lidar: experimental study using a 1D+4D-Var technique. *Quarterly Journal of the Royal Meteorological Society* 2015; 141: 2708-2725. doi: 10.1002/QJ.2558
19. Fielding MD, Janisková M. Direct 4D-Var assimilation of space-borne cloud radar reflectivity and lidar backscatter. Part I: Observation operator and implementation. *Quarterly Journal of the Royal Meteorological Society* 2020; 146: 3877-3899. doi: 10.1002/QJ.3878
20. Janisková M, Fielding MD. Direct 4D-Var assimilation of space-borne cloud radar and lidar observations. Part II: Impact on analysis and subsequent forecast. *Quarterly Journal of the Royal Meteorological Society* 2020; 146: 3900-3916. doi: 10.1002/QJ.3879
21. Cherukuru NW, Calhoun R, Krishnamurthy R, Benny S, Reuder J, Flügge M. 2D VAR single Doppler lidar vector retrieval and its application in offshore wind energy. *Energy Procedia* 2017; 137: 497-504. doi: 10.1016/j.egypro.2017.10.378

22. Cherukuru NW. *Doppler Lidar Vector Retrievals and Atmospheric Data Visualization in Mixed/Augmented Reality*. PhD thesis. Arizona State University, 1151 S Forest Ave Tempe, AZ 85281, USA; 2017.
23. Technariumas . Inpainting Healing algorithm. <https://github.com/Technariumas/Inpainting;> . Accessed: 2022-09-22.
24. Clifton A, Boquet M, Burin Des Roziers E, et al. Remote Sensing of Complex Flows by Doppler Wind Lidar: Issues and Preliminary Recommendations. Tech. Rep. December, National Renewable Energy Laboratory (NREL); Golden, CO (United States): 2015

How to cite this article: Krutova M., Bakhoday-Paskyabi, M., and Reuder J. (2023), Validation of the 2D-VAR lidar retrieval algorithm for non-homogeneous wind fields, *Wind Energy*, .