Wavelet-based optical flow for two-component wind field estimation from single aerosol lidar data

Pierre Dérian, Christopher F. Mauzey and Shane D. Mayor*

California State University Chico, Chico, California

**Corresponding author address:* California State University Chico, 400 W. First St., Chico, CA 95928

E-mail: sdmayor@csuchico.edu

ABSTRACT

A motion estimation algorithm is applied to image sequences produced by a horizontally-scanning elastic backscatter lidar. The algorithm, a wavelet-based optical flow estimator named *Typhoon*, produces dense two-component vector flow fields that correspond to the apparent motion of microscale aerosol features. To validate the efficacy of this approach for the remote measurement of wind fields in the lower atmosphere, an experiment was conducted in Chico, California, in 2013 and 2014. The flow fields, estimated every 17 s, were compared with measurements from an independent Doppler lidar. Time-series of wind speed and direction, statistical assessment of the 10-min averages and examples of wind fields are presented. The comparison of 10-min averages at 100 m AGL reveals excellent correlations between estimates from the *Typhoon* algorithm and measurements from the Doppler lidar. Power spectra and spectral transfer functions are computed to characterize the filtering effects of the algorithm in the spatial domain.

1. Introduction

Motion estimation is a branch in the field of computer vision that develops algorithms to determine the apparent movement of objects in sequences of digital images. Since the seminal paper by Horn and Schunck (1981), the applications of these numerical methods have become numerous; they play key roles in the success of many modern technologies including bioinformatics, video compression and machine vision. These techniques are also commonly found in experimental fluid dynamics, applied for example to particle image velocimetry (PIV) (Adrian 2005). In contrast to in-situ measurements which are inherently restricted to a single point of space, motion estimation methods are non-intrusive and provide fields or volumes of velocity vectors and thus offer a broader perspective of the flow.

Because of the abundance of images in the atmospheric and oceanic sciences, motion estimation has been practiced since before the digital age. For example, determination of the movement of cloud or water vapor features in satellite images was done prior to the work of Horn and Schunck (1981) through a block-matching approach (Leese et al. 1971). These atmospheric motion vectors (AMV) constitute nowadays an essential component of the observations assimilated by numerical weather prediction models (García-Pereda and Borde 2014). Other modern applications involve for example the recovery of glacier velocities (Scambos et al. 1992), displacements resulting from landslides (Stumpf et al. 2013), surface water flows (Dugan et al. 2014) and breaking waves dynamics (Melville and Matusov 2002).

Another application, similar to PIV and AMV, involves the estimation of 2D, 2-component wind field from the apparent motion in aerosol backscatter lidar data (Schols and Eloranta 1992). Thus far the motion estimation algorithms used in that context were variations of the cross-correlation method (Mayor et al. 2012; Hamada et al. 2015). In this paper, a more recent approach that was devised specifically for application to fluid motion is investigated. This algorithm, named *Typhoon*, is a wavelet-based optical flow estimator. It was previously validated with synthetic and real PIV images (Dérian 2012). Here, as a first step, the validity of this wavelet-based optical flow approach in the context of atmospheric lidar data is demonstrated.

The paper is organized as follows: Section 2 introduces the motion estimation framework for the wind measurement problem and the traditional cross-correlation algorithm. Section 3 presents the proposed *Typhoon* algorithm. The input aerosol backscatter lidar data is detailed in Section 4. Finally, in Section 5, estimated wind fields are validated by comparisons with remote measurements from a commercial Doppler lidar. Power spectra and transfer functions are calculated to show the filtering effect of the proposed approach.

2. Wind measurement and motion estimation

a. Wind measurement strategies

Air motion is represented by a three-component vector and may be defined at all points in the atmosphere. The wind is generally regarded as the vector consisting of two horizontal components. Active remote wind measurement techniques may be subdivided into Doppler and non-Doppler approaches.

Ground-based radars and lidars typically collect data in a spherical coordinate system. Doppler radars and lidars directly measure only the radial (line-of-sight) component of air motion. For a Doppler radar or lidar to measure the wind, specific scanning strategies and assumptions about the air motion over space and time must be made. Wind profiling describes the use of a remote sensor to provide a vertical profile of horizontal wind vectors at a single location above the surface of the earth. Alternatively, two Doppler radars or lidars, separated by some horizontal distance, may be used to probe an area from different angles and obtain a two-component wind field. This approach is known as "dual-Doppler" (Stawiarski et al. 2013).

Non-Doppler approaches estimate wind fields from the spatial and temporal movement of features observed by the instrument. Eloranta et al. (1975) provided some of the first remote wind measurements by lidar in the lower atmosphere. Since that time, hardware and software has advanced greatly and a small number of validation experiments have been conducted, e.g. Mayor et al. (2012). Meanwhile, other fields, in particular experimental fluid dynamics, have developed similar approaches to retrieve motions. This concept is also known to the computer vision community, where it is associated with the wide family of *motion estimation* techniques.

b. Fluid motion estimation: the vision approach

The idea of using the *apparent motion of tracers* to infer the invisible underlying fluid flow is not new. It "could probably be traced far back in history to the first time a person possessing the concept of velocity watched small debris moving on the surface of a flowing stream" (Adrian 2005). Many visualization methods have been developed, such as using droplets, dye, smoke or shadows for the purpose of revealing fluid flow structures and dynamics (Van Dyke 1982). This led in particular to the well-known PIV techniques, which have been used in experimental fluid dynamics for almost 30 years (Adrian 2005). Our 2D, 2-component wind measurement approach fits in the motion estimation context: the tracers are the aerosol *features*, visualized by the lidar system, and the motion estimation technique is usually the cross-correlation. This configuration is very comparable to PIV, with the important differences that the distribution of aerosols in the atmosphere (the "seeding" of the flow) cannot be controlled, and that the images are not of individual particles, but instead of a field that approximately represents particle concentration (Held et al. 2012). In these aspects, this problem is closer to AMV computation. An important difference is the temporal and spatial resolutions covered by these two approaches: typically, on the order of 15 seconds and 10 meters for the considered lidar data, versus 15 minutes and several kilometers for geostationary satellite imagery (García-Pereda and Borde 2014).

c. Motion estimation framework

Motion estimation aims to recover the apparent displacements within a sequence of images. The time and space variations of an observable image quantity are used to infer the underlying motion field occurring in the image plane between two consecutive frames of the sequence. In this work, input images are the scans provided by the lidar, the movements of the variations of aerosol backscatter intensity are used to estimate the wind field.

In the following, the scan domain is noted as $\Omega \subset \mathbb{R}^2$. The observable backscatter intensity is noted as $I_n(\mathbf{x})$ at pixel $\mathbf{x} = (x_1, x_2) \in \Omega$ and at discrete time t_n, $n \in \mathbb{N}$. The *apparent displacement* between two consecutive scans I_n, I_{n+1} is a 2D vector field $\mathbf{u}$:

$$\mathbf{u}(\mathbf{x}, t_n) = \begin{pmatrix} u_1(\mathbf{x}, t_n) \\ u_2(\mathbf{x}, t_n) \end{pmatrix}.$$

This displacement is measured in pixel units and occurs over the time $\delta t_n = t_{n+1} - t_n$ s. If the scan has a resolution of δx m pixel^{-1}, an estimation of the *instantaneous wind velocity* $\mathbf{v}$ in m s^{-1} is therefore given by:

$$\mathbf{v}(\mathbf{x}, t_n) = \frac{\delta x}{\delta t_n} \mathbf{u}(\mathbf{x}, t_n). \tag{1}$$

As such, the motion is assumed to be stationary during the time step δ_t.

Velocity components v_1, v_2 are the *in-plane* components, that is, they belong to the image plane. Due to the very low value of the elevation angle of the lidar scan plane (typically $< 6°$), these components coincide with the horizontal wind components (usually denoted u, v in atmospheric sciences). The *out-of plane* component (normal to the scan plane), which remains unestimated, thus corresponds to the vertical component w.

The question of the accuracy of motion estimation techniques is often raised. The answer is complex, since it involves the *data characteristics* (spatial, temporal resolutions), the information given by the *visualization method* (the image content), and the underlying *motion field* itself. In the current context, the later contributions are difficult to quantify, as they depend largely on the conditions (e.g., the presence of particulate matter, the scales and variability of the wind field). However, assuming ideal conditions and a perfect model, errors related to the resolution of data may be quantified. If displacements are measured as integers on the image grid, the systematic error is ± 0.5 pixel, which then gives $\pm 0.5\delta_x/\delta_t$ m s^{-1} for each motion component. In practice, various interpolation techniques allow for sub-pixel estimation, reducing this error. The error can be also lowered by using a smaller δ_x and/or a larger δ_t. However, for a given motion field, a smaller δ_x results in larger apparent displacements, which can be more challenging for estimation algorithms. On the opposite, larger δ_t leads to less accurate perception of the instantaneous velocity, since the assumption of stationarity of the motion field is less valid over longer periods.

Any motion estimation technique features two main aspects. The first one, known as the *data model*, describes the link between observations I (the aerosol backscatter intensity) and the underlying unknown displacement $\mathbf{u}$. This model should take into account the nature of observed data and its relevant dynamics. Then, as an inverse problem, motion estimation is usually ill-posed. The second aspect is therefore the *regularization*, which is required in order to close the estimation problem. The regularization may also provides information where the data model fails locally. The various estimation techniques feature different data models, regularizations or implementation strategies.

d. The cross-correlation algorithm, concept and limitations

The cross-correlation technique performs independent, local motion estimations on subregions (blocks) of the scan domain. It consists in correlating a block of the first scan I_n with a translated block of the second scan I_{n+1}; the translation vector $\mathbf{u}$ which induces a correlation peak is considered to be the displacement at the center of the block (Schols and Eloranta 1992). The estimation problem, presented in its basic form, is written as:

$$\forall \mathbf{x} \in \Omega_C, \mathbf{u}(\mathbf{x}, t_n) = \arg\max_{\mathbf{u}} \sum_{\mathbf{y} \in B(\mathbf{x})} \frac{[I_{n+1}(\mathbf{y}+\mathbf{u}) - \mu_{n+1}(\mathbf{x}+\mathbf{u})]\,[I_n(\mathbf{y}) - \mu_n(\mathbf{x})]}{\sigma_{n+1}^2(\mathbf{x}+\mathbf{u})\sigma_n^2(\mathbf{x})}, \tag{2}$$

where $\Omega_C \subset \Omega$ is the set of block centers (and therefore the set of locations of estimated vectors), $B(\mathbf{x})$ is the block centered on $\mathbf{x}$, $\mu_p(\mathbf{x})$ and $\sigma_p(\mathbf{x})$ are the mean and standard deviation, respectively, of backscatter intensity I_p over block $B(\mathbf{x})$. Note that in practice, this cross-correlation function (CCF) is computed using the FFT for computational efficiency.

In this case, the data model is the CCF (2) itself; the regularization is implicitly given by the size of block $B(\mathbf{x})$ which should be large enough to contain reliable information, yet as small as possible to resolve small scale motions. Typically, neighboring blocks overlap by 50%, so that the estimated motion field is *sparse* (fewer motion vectors than pixels). Each vector is the result of a single *independent* problem, which makes the CCF algorithm pleasingly parallel (Mauzey et al. 2012). This cross-correlation approach and its numerous variants have become widely used in PIV (Adrian and Westerweel 2010); in geosciences it is often applied to satellite imagery to retrieve for instance glacier velocities (Scambos et al. 1992). It is also the standard method to derive AMVs (Schmetz et al. 1993; García-Pereda and Borde 2014), and has given good results with aerosol backscatter lidar data, as shown in Schols and Eloranta (1992), Mayor and Eloranta (2001) and Mayor et al. (2012).

However, this method as presented in (2) is not exempt from drawbacks. First, the displacement within an entire block $B(\mathbf{x})$ is explained by a single vector $\mathbf{u}(\mathbf{x})$, which implies that this displacement is assumed to be uniform (constant) over the block. The larger the block, the less likely this assumption is to be true. Yet, as overly small blocks may result in uncertainties due to lack of information, "large" blocks are usually preferred. This leads to the second point: as displacements occurring within large blocks are likely not uniform, the estimated $\mathbf{u}(\mathbf{x})$ corresponds to a power-weighted average of the apparent displacements within the corresponding block $B(\mathbf{x})$ (Hamada 2014), which results in an over-smoothed motion field. To address these issues, this study proposes to evaluate a recently developed motion estimation algorithm dedicated to fluid flows.

3. Typhoon algorithm

Early attempts with a different class of motion estimation methods, often called *optical flow*, were conducted in 2010 on the CHATS[1] dataset and led to promising results (Dérian et al. 2010). Since then the authors developed a new version of the algorithm based on a *wavelet* framework, named *Typhoon*. The extensive description of the algorithm is largely mathematical and details regarding the design of the data-model and the regularization can be found in Dérian et al. (2013) and Kadri Harouna et al. (2013), respectively. In the following, an overview of the method and the improvements made to achieve real-time wind estimation from aerosol backscatter lidar imagery are provided.

[1] Canopy Horizontal Array Turbulence Study, near Dixon, CA, 2007 – see Patton et al. (2011).

a. Optical flow, from observations to motion

The proposed approach has two major differences with respect to the cross-correlation algorithm presented above. First, this wavelet-based optical flow uses a *global* formulation: all vectors $\mathbf{u}(\mathbf{x})$ of the displacement field $\mathbf{u}$ are estimated simultaneously by solving a single problem, whereas the cross-correlation approach in (2) has as many independent problems as vectors $\mathbf{u}(\mathbf{x})$. Second, this method provides a *dense* estimate, that is to say one displacement vector at every point $\mathbf{x}$ of the scan domain Ω, whereas the CCF solution is usually sparse. The estimate is obtained by minimizing a functional, similar to an energy, defined over the whole scan domain:

$$\mathbf{u} = \arg\min_{\mathbf{u}} \left\{ \frac{1}{2} \int_{\Omega} [f_{data}(I, \mathbf{u})]^2 \, d\mathbf{x} + \frac{\alpha}{2} \int_{\Omega} [f_{reg}(\mathbf{u})]^2 \, d\mathbf{x} \right\}. \tag{3}$$

f_{data} is the data model that depends on observations I and unknown displacement $\mathbf{u}$, while the regularization f_{reg} depends on $\mathbf{u}$ only. The parameter $\alpha > 0$ balances the two terms and is fixed by the user.

The data model used in *Typhoon* is known as the displaced frame difference (DFD):

$$I_{n+1}(\mathbf{x} + \mathbf{u}(\mathbf{x}, t_n)) = I_n(\mathbf{x}). \tag{4}$$

It is analogous to finding the displacement field $\mathbf{u}$ that "warps" an image into the next one. This model assumes the *consistency of backscatter intensity* along the trajectory of an aerosol feature during the time interval $[t_n; t_{n+1}]$, that is to say an aerosol feature will present the same intensity, the same "signature", in both scans I_n, I_{n+1}. Therefore any phenomena inducing a significant change in intensity, such as turbulent diffusion or out-of-plane motion, can possibly lead to *false apparent motions*.[2] Such phenomena are not uncommon, but it can be reasonably assumed that the time scales at which they act are significantly larger than the inter-scan time-step δt_n, so that

[2]False apparent motions refer here to illusory motions of aerosol features that do not correspond to the horizontal wind.

the DFD (4) remains valid. It is also important to note that from formulation (3), the data model is not *strictly* enforced. Instead, the solution achieves a balance between trying to follow the model on one hand and the regularization on the other – hence the role of the parameter α, which allows the user to give more weight to one term over the other.

Regularization schemes usually encourage the estimate $\mathbf{u}$ to follow some *smoothness assumption*. This work uses the most simple first-order regularization, originally introduced in Horn and Schunck (1981), which penalizes strong velocity gradients. For each displacement component u_i, $i = 1, 2$:

$$f_{reg}(u_i) = |\nabla u_i| = \sqrt{\left(\frac{\partial u_i}{\partial x_1}\right)^2 + \left(\frac{\partial u_i}{\partial x_2}\right)^2}. \tag{5}$$

Note that the square root is later cancelled by the square in (3). If the regularization is given much more weight than the data model ($\alpha \to \infty$ in (3)), the solution that minimizes (3) moves toward a uniform motion field (with $\nabla u_i = 0$ for $i = 1, 2$). The regularizer also takes precedence over the data model locally where the latter is inefficient, for instance within uniform regions of the input images. Other regularizers are available in *Typhoon*, penalizing, for instance, the vorticity or divergence of the flow, or the gradient of vorticity, divergence; some of these schemes have proven to be very efficient with PIV and water vapor satellite images (Corpetti et al. 2002). However, as the regularization becomes more complex, the associated computational costs increase, which may reduce the ability to achieve real-time estimation. Moreover, in the context of aerosol backscatter lidar images, little to no improvement brought by the use of these advanced schemes was found. This could be linked to the specificities of this lidar data, which will be detailed further in Section 4.

The DFD model (4) and the Horn and Schunck regularizer (5) inserted into (3) complete the motion estimation problem:

$$\mathbf{u}(t_n) = \arg\min_{\mathbf{u}} \left\{ \frac{1}{2} \int_\Omega \left[I_{n+1}(\mathbf{x} + \mathbf{u}(\mathbf{x}, t_n)) - I_n(\mathbf{x}) \right]^2 d\mathbf{x} \right. \\ \left. + \frac{\alpha}{2} \int_\Omega \sum_{i=1,2} |\nabla u_i(\mathbf{x}, t_n)|^2 \, d\mathbf{x} \right\}. \tag{6}$$

A particularity of this problem is that the DFD model (4) is not linear in $\mathbf{u}$, so that the whole functional is not quadratic. This complicates the minimization process, as the existence of a global minimum is not guaranteed. This is another role for the regularization term: it convexifies the functional as $\alpha \to \infty$. But, as large α values are unmanageable, to ensure a successful minimization it is important for the solution $\mathbf{u}$ to lie "close" to the first guess.[3] This calls for the use of an incremental strategy, often known as "*multi-resolution*": the displacement field is estimated following a coarse-to-fine process, starting with coarse structures of large amplitudes, and progressively refining toward smaller scales. This last point motivates the use of the *wavelet framework*.

b. Introduction to the wavelet framework

In signal processing, the spectral space is often used to analyze or exhibit some properties of a given signal. The FFT leads to a representation in terms of sine and cosine functions of specific frequencies. Any spatial information is lost in the process: the Fourier coefficients, which form an equivalent representation of the input signal, yield no information as to *where* their associated frequency is or is not present. This is due to the fact that the sine and cosine functions, which form the basis of the spectral space, are very well localized in frequency but have an infinite support in space. Conversely, looking at the signal in the physical space does not give any information on the frequency content. The wavelet formalism offers a trade-off: the wavelet functions are localized both in space and frequency, thus they enable access to information on the *frequency content and*

[3]which is usually the null motion field, $\mathbf{u}(\mathbf{x}) = 0 \; \forall \mathbf{x} \in \Omega$.

the spatial location simultaneously – at the cost of lower precision. A wavelet representation of a given signal consists of a *coarse approximation* of the signal, along with several sets of *details* containing spatially-localized information at various ranges of frequencies. Note that instead of frequency, the wavelet formalism prefers the equivalent but reciprocal notion of *scale*.

This multi-scale (or, multi-resolution) representation offered by the wavelet transform is the main motivation to adopt wavelet bases for displacement components u_1, u_2. It leads to a "natural" coarse-to-fine strategy suitable to motion estimation (Dérian et al. 2013). Approximation and coarse detail coefficients are estimated first, then fine-scale details are successively added until the smallest scale is reached. Besides the multi-scale framework, wavelet bases also allow the representation of arbitrary regular functions (a 3D fluid motion field should at least be continuous). While the continuity might not be a relevant assumption for the 2D field, a sufficient regularity is required in order to compute the regularizing terms presented in Section 3.a, which involve spatial derivatives. Finally, these regularization schemes find a relatively simple yet very accurate implementation in that context (Kadri Harouna et al. 2013). Similarly to the Fourier transform, the wavelet transform is a linear, separable[4] operator, with fast algorithms (fast wavelet transform, FWT) for computational efficiency. Wavelets are also used in many fields, from signal denoising to video compression; Mallat (2008) discusses an extensive presentation of the theory and applications.

Conceptually, the use of wavelet bases does not lead to significant changes to the estimation problem (6). Each motion component u_i is expressed as the inverse transform (reconstruction) of its corresponding wavelet coefficients c_i:

$$u_i = W_{inv}(c_i)\,, \quad i = 1,2\,,$$

[4]The 2D transform is obtained by combining two 1D transform, first along rows then along columns.

where W_{inv} denotes the inverse wavelet transform. The set of wavelet coefficients $\{c_1, c_2\}$ thus is the unknown to the estimation problem.

c. Recent improvements

The original algorithm detailed in Dérian et al. (2013) would accept square images only. If input images were rectangular, they had to be padded to turn them square, which increases the computational burden. The current version has been modified to accept rectangular images.

The main improvement is the result of redesigning the code to run in "real-time". To keep up with real-time, the estimate of wind field $\mathbf{v}(t_n)$ from scans I_n, I_{n+1} must be complete by the time the next scan I_{n+2} is made available, with the inter-scan time-step δt_n typically on the order of 10 to 20 seconds. Since the whole motion field is estimated simultaneously, the number of variables is quite large: a dense estimate from 512×512 pixel images represents about half a million unknowns. Wavelet transforms lie at the core of the estimation process. Each evaluation of the functional (6) requires two inverse FWTs (to reconstruct the displacement $\mathbf{u}$ from its coefficients) and two forward FWTs (to compute the gradient). In order to achieve the necessary reduction in computation time, the low-level functions of the algorithm – in particular, the wavelet transforms – were rewritten in *CUDA language*, which enables it to execute on NVIDIA's graphic processing units (GPU). GPUs designed for scientific computing rely on several thousands of small computing units, thus providing massive parallelization capabilities. The CUDA version of *Typhoon* running on an NVIDIA GeForce GTX Titan is 10 to 100 times faster than the original version (Mauzey et al. 2014), and is sufficient to meet the real-time requirements.

4. Application to aerosol backscatter data

The results presented hereafter have been obtained from data collected by the Raman-shifted Eye-safe Aerosol Lidar (REAL) (Mayor and Spuler 2004; Spuler and Mayor 2005; Mayor et al. 2007; Spuler and Mayor 2005) in 2013 and 2014 in Chico, California. The REAL is a ground-based, scanning, elastic backscatter lidar operating at a wavelength of 1.54 microns, with a pulse energy typically between 120 and 170 mJ, a pulse rate of 10 Hz and a pulse duration of 6 ns. It employs 40 cm diameter optics and an analog direct detection receiver. The backscatter signal is sufficiently strong from a single pulse that averaging over multiple pulses is not required. This section describes the input scan data as well as the preprocessing steps.

a. Data preprocessing

Before motion estimation takes place, the raw signal delivered by the REAL must be preprocessed. Lidar data is sampled on a polar grid, with the lidar at the origin. Each scan is composed of *shots*, with a shot being a 1D array of backscatter samples, uniformly spaced along the range r every 1.5 m, collected at a given angular position θ from a single laser pulse.

The raw backscatter intensity $I_{raw}(r,\theta)$, with the range r and the azimuth angle θ, corresponds to the actual backscatter signal $\beta(r,\theta)$ and an additive noise $\varepsilon(r,\theta)$.

$$I_{raw}(r,\theta) = \beta(r,\theta) + \varepsilon(r,\theta)\,.$$

The noise ε combines contributions from the atmosphere and the instrument and can be modeled by a random variable which follows a normal distribution of mean μ_θ and standard deviation σ_θ. Values of μ_θ, σ_θ change slightly from one shot to another, hence their dependency in θ; they can be estimated for each shot from background data. As explained in Mayor et al. (2012), first the noise mean is subtracted:

$$I_0(r,\theta) = I_{raw}(r,\theta) - \mu_\theta = \beta(r,\theta) + \varepsilon_0(r,\theta)\,,$$

with $\varepsilon_0(r,\theta) = \varepsilon(r,\theta) - \mu_\theta$ the now centered random noise. The raw signal-to-noise ratio (SNR) is computed at that point:

$$SNR_{raw}(r,\theta) = \frac{I_0(r,\theta)}{\sigma_\theta}. \tag{7}$$

Shots are then multiplied by the square of the range to compensate for the one-over-range-squared decay of the backscatter β:

$$I_{r^2}(r,\theta) = r^2 I_0(r,\theta) = r^2\beta(r,\theta) + r^2\varepsilon_0(\theta)\,.$$

Note that the noise amplitude now increases as the square of the range. For optimal results, it is then essential to discard irrelevant noisy data, which is discussed further.

After conversion to decibels, shots are filtered in the range dimension. The low-pass median filter of length 7 points (10.5 m) removes high-intensity spikes typically caused by hard-targets such as birds and insects, while the high-pass median filter of length 333 points (500 m) removes the very large structures to reveal local fluctuations. Figure 1 presents an example of preprocessed backscatter data (panel a), along with the corresponding raw SNR (7) (panel b).

b. Detecting coherent features

Two different aspects complicate the motion estimation process. First, due to the nature of backscatter data, the raw SNR (7) decays as one-over-range-squared. Typically, for the REAL operating in Chico, CA, the SNR resulting from a single laser pulse drops below 5 at $r = 3$ km. Such high levels of noise in the far range are challenging for optical flow. Second, for the purpose of motion estimation, a good SNR in the near range does not necessarily imply useful information.

For instance, coherent features can be absent from a region of the scan, yielding much uncertainties as to the underlying wind field in that region.

In order to maximize the quality of the results, the scan areas presenting no coherent aerosol features are discarded. Because of the regularization schemes provided by optical flow (Section 3.a), wind vectors estimated over noisy areas *could* be relevant. However judging so proves to be difficult, as often even a basic visual confirmation is impossible in noisy regions. Hence, it is safer to simply discard the noisy image data before motion estimation.

To detect the presence of coherent aerosol features, the *image SNR* is used. It is defined as the ratio of the *local* standard deviation of coherent signal $\sigma_\beta(r,\theta)$ to the *local* standard deviation of noise σ_ε:

$$SNR_{img}(r,\theta) = \frac{\sigma_\beta(r,\theta)}{\sigma_\varepsilon(r,\theta)}. \tag{8}$$

This ratio is estimated from the autocovariance function of preprocessed data $I(r,\theta)$. For every point (r,θ), the autocovariance C_l is computed along the range from data in $[r-l/2; r+l/2]$. Then, the local variance of coherent signal is given by the average of coefficients at lag 1 and -1:

$$(\sigma_\beta)^2 = 0.5\,(C_l(-1)+C_l(1)),$$

while the local variance of noise is obtained from the 0-lag coefficient and σ_β:

$$(\sigma_\varepsilon)^2 = C_l(0) - (\sigma_\beta)^2.$$

An example of image SNR is shown in Fig. 1 panel c. A 256-point window was used to compute the autocovariance, corresponding to $l = 384$ m.

From the image SNR, a valid data domain is computed for each scan. It is assumed that the best data is in the near range, therefore the valid domain is simply defined by a far-range boundary. For each shot (azimuth θ), this far-range boundary is given by the smallest range $R(\theta)$ above which the image SNR remains below a threshold τ fixed by the user:

$$\forall\theta,\ R(\theta) = \min_{R}\left\{R:\ \forall r > R,\ SNR_{img}(r,\theta) < \tau\right\}\ . \tag{9}$$

Finally, a low-pass median filter of width 25 points and a Gaussian filter of parameter $\sigma = 2$ points are applied to the set of $R(\theta)$, to exclude small isolated features and smooth the boundary. An example of mask representing the valid data domain is shown in Fig. 1 panel d, using $\tau = 3$.

c. Correction of image distortions

A lidar scan does not correspond to an instantaneous view of the aerosol distribution. The shots that compose the scan are acquired sequentially. In the event of high wind speeds, this leads to apparent distortions of the aerosol features in the lidar images, which in turn causes the estimated motion to be biased. This issue was first noted by Sasano et al. (1982) who proposed an iterative correction method. Assuming that the aerosol features are transported *without deformation* by a uniform wind vector, scans can be warped to reconstruct an approximated instantaneous view of the aerosols, thus improving the accuracy of motion estimation. In this study, implementation proceeds as follows for a given scan pair:

i. Estimate the displacement field $\mathbf{u}$ from the pair of scans with the *Typhoon* algorithm.

ii. Convert to velocity field $\mathbf{v}$ using (1).

iii. Correct both scans for distortions using wind field $\mathbf{v}$, following Sasano et al. (1982). The time of the beam at the center of the scan is used as the reference time.

iv. Repeat i–iii until mean wind speed $|\bar{\mathbf{v}}|$ changes by either less than 1%, or less than $0.25\delta x/\delta t$. Typically, it requires 2-3 iterations.

The correction step (iii) is carried on the polar grid data. After correction, backscatter data is no longer known on a regular polar grid, but instead is at scattered locations.

d. Cartesian gridding

After preprocessing, masking and correction for distortions, the backscatter data is interpolated on a Cartesian grid of spacing $\delta x = 8$ m. It is possible to perform the motion estimation directly on the original polar grid, however, as mentioned above, the correction step destroys the regularity of the mesh. Fast interpolation on large sets of scattered data can be challenging, considering real-time requirements. In this work, a CUDA implementation of nearest-neighbor interpolation was used.

5. Validation

A field experiment was conducted in Chico, CA, from mid-September 2013 to mid-January 2014, to validate the wind fields recovered by *Typhoon*. A Doppler lidar (DL) was deployed to provide independent wind measurements. It is a pulsed, heterodyne detection Doppler lidar commercialized by HALO Photonics under the name Stream Line (Pearson et al. 2009). This model was previously certified against cup anemometer measurements (Axel and Ailt-Wiard 2014), and showed very good agreement to radar wind profiler and radiosonde (Päschke et al. 2014). The DL has the S/N 0811-35 and was built in November 2011. It operates at a wavelength of 1.5 microns, pulse energy of 20 μJ, pulse rate of 15 kHz, and a pulse duration of 150 ns. DL data was filtered following the manufacturer's indications, keeping only points for which the minimum SNR intensity > 1.01.

Since it is not possible to retrieve a 2D 2-component wind field using a single DL, two different configurations were investigated.

- *Temporal validation*. The DL was located at 1500 m range, 15° azimuth from the REAL and operated in vertical profiling mode. Data from this configuration enable comparisons of time-series of *2-component* wind velocities *at the DL location*. This phase of the experiment was conducted during September and October 2013.

- *Spatial validation*. The DL was located on the roof of the REAL container and operated in fixed-beam mode, staring at the center of the sector scan area swept by the REAL. This configuration enables one to compare *radial wind* velocity components *along the DL line-of-sight*. Data for this second phase of the experiment were collected in December 2013 and January 2014.

The main parameters used by both systems during these two experiments are summarized in Table 1.

a. Temporal validation

In this experiment, the REAL scans between -15° and 45° azimuth, with a 4° elevation, every 17 s. This places the scan at 100 m AGL at the range of the DL. The DL operates in vertical profiling mode (VAD scan), providing a profile of 2-component horizontal wind vector about every 15 s.

A typical example of aerosol motion estimation is presented in Fig. 2. It features a close-up of two motion fields estimated from three successive position plan indicator (PPI) scans. The flow is relatively uniform, and can be visually identified due to a large aerosol feature that moves toward the southeast. The DL wind vectors at 100 m AGL are displayed for comparison and show a good agreement with the *Typhoon* estimates.

In this paper, an effort is made to establish the potential of the *Typhoon* algorithm when applied to aerosol backscatter lidar data. However, quality of the data depends upon the performance of the instrument and the state of the atmosphere. Therefore, we selected the days presenting the best potential for this validation among data collected in Chico, CA from mid-September to mid-November 2013, with the expectation that future advances in hardware will lead to increases in data quality and availability. First, due to the local typical conditions in Chico, aerosol backscatter imagery is much better for this application during the daytime than during nighttime. Therefore, this study was restricted to daytime only. Second, the percentage of valid backscatter data (Sec. 4.b), during daytime, in a 50 m radius around the DL were computed. These values are plotted against the mean wind speed measured by the DL the same day in Fig. 3. With a sufficient spatial distribution of aerosol features, dense 2-component wind fields can be delivered up to several km in range. Figure 4 shows an example of such wind field on a day with high speed and uniform direction, with vectors available out to 4 km range. The low-SNR area in the far-range were dynamically excluded. Figure 5 presents a view of a ≈ 200 m vortex, illustrating the ability of *Typhoon* to extract coherent structures at intermediate scales.

Three specific cases are described below: light, moderate and strong wind conditions. These days are represented by solid diamonds in Fig. 3. For each case, time-series of instantaneous and 10-min averaged wind measurements are presented. 10-min averages are the reference measures for instrument validation in the wind power industry (Bailey 2012). Then, statistics on 10-min averages for the 15 days having more than 85% valid data are presented.

The VAD scan strategy used by the DL assumes that the wind is uniform throughout the swept area (Mann et al. 2009, 2010; Sathe et al. 2011; Sathe and Mann 2012); in this case this region is a disc of about 100 m radius, represented by a turquoise circle in Figs. 4 and 5. In order to compare results of the study to the DL measures, instantaneous *Typhoon* estimates are averaged in space over a similar sized area centered on the DL location.

Occasionally, the estimation may fail and result in obvious outliers. Those outliers can be detected and removed under the assumption of temporal coherence of the wind field. The *normalized* *median test*, commonly used in PIV (Adrian and Westerweel 2010), was implemented. Similar concepts are used with radar wind profilers (Weber et al. 1993). Within each 10-min window, the median wind vector $\mathbf{v_m}$ is computed, as well as the residuals $r(\mathbf{v}) = |\mathbf{v_m} - \mathbf{v}|$ for each vector $\mathbf{v}$ of the window. Vectors for which the residual $r(\mathbf{v})$ is twice larger than the median of residuals r_m are discarded.

1) LIGHT WIND CASE

Figure 6 shows wind speed and direction measured by the DL at 100 m AGL and estimated by *Typhoon* for a 12-hour period starting on October 23 at 15:00 UTC. It is a light wind episode with speeds remaining below 3 m s^{-1} and variable direction. Estimates are missing over a period approximatively covering 15:00 to 17:00 UTC. This is due to the coherent feature detection presented in Sec. 4.b: no significant features were present in the region of interest at that time, therefore no motion estimates are available. Then, between 17:00 and 18:00 UTC, *Typhoon* speed and direction estimates are in systematic error. Visual inspection of the aerosol imagery reveals the mixed layer growing with the entrainment zone passing through the altitude of the intercomparison. It appears that the plumes and wind shear in the entrainment zone result in false apparent motions that bias the motion estimations. Later, two reversals of wind direction occurred at 22:30 and 23:30 UTC that correspond to the passage of a vortex of diameter ≈ 200 m over the region of interest (see also Fig. 5 for a spatial visualization). This microscale circulation resembles those that have resulted from large eddy simulation of convective boundary layers (Schmidt and Schumann 1989; Kanak 2005; Sullivan and Patton 2011). Correlation coefficients R^2 for the 10-min averaged wind components are 0.951 and 0.600 for u and v, respectively. Excluding the 17:00 – 18:00 UTC period with false apparent motions, R^2 values increase to 0.966 and 0.866.

2) Moderate wind case

Figure 7 shows wind speed and direction measured by the DL at 100 m AGL and estimated by *Typhoon* for a 12-hour period starting on September 17 at 15:00 UTC. This wind episode features speeds ranging 0 to 10 m s^{-1} and direction mostly stationary except for a 2-hour fluctuating episode (corresponding to the lowest wind speeds). Wind speed is underestimated at two occasions, both corresponding to rapid and large changes in direction around 22:30 and 23:00 UTC. Otherwise, both series of data are in very good agreement. This is confirmed by the 10-min averaged wind components: correlation coefficients R^2 are 0.979 and 0.991 for u and v, respectively.

3) Strong wind case

Figure 8 shows wind speed and direction measured by the DL at 100 m AGL and estimated by *Typhoon* for a 12-hour period starting on October 9 at 15:00 UTC. It is a strong wind episode with speeds up to 16 m s^{-1} and very consistent flow from the northwest direction. Both time-series are again in very good agreement. Correlation coefficients R^2 for the 10-min averaged wind components are 0.984 and 0.929 for u and v, respectively.

4) Overall considerations

Scatter plots of 10-min averaged wind components measured during the daytime[5] for the 15 "best" days (Fig. 3) are presented in Fig. 9. They show an overall excellent agreement of *Typhoon* estimates with DL measurements at 100 m AGL: correlation coefficients R^2 are 0.995 and 0.997 for u and v, respectively. Detailed statistics on u and v are available in Tables 2 and 3. In terms of wind speed, a linear regression gives a slope of 1.000 with an offset of -0.10 m s^{-1}, R^2 coefficient

[5]"Daytime" is arbitrarily considered to be 15:00 – 01:00 UTC (10 hours).

is 0.991. Regarding the wind direction, the offset is 1.1° and R^2 coefficient is 0.944.[6] This $\approx 1°$ offset observed for the direction corresponds to the precision at which the DL was oriented during its deployment. The root mean square error (RMSE) between *Typhoon*'s estimates and the DL observations is 0.29 m s^{-1} on both on u and v components. This is slightly higher than the expected systematic error of $0.5\delta_x/\delta_t \approx 0.24$ m s^{-1} which assumes perfect data and model (Sec. 2.c). The few remaining outliers mostly correspond to false apparent motions, typically occurring at the beginning and end of the day as the boundary layer depth evolves.

From the time-series shown in Figs. 6, 7 and 8, it appears the variability of the wind speed obtained by *Typhoon* is less than that measured by the Doppler. Figure 10 is a scatter plot of turbulent kinetic energy (TKE) as measured by the Doppler and *Typhoon* over 10-min intervals. A linear regression suggests that the TKE from *Typhoon* is about 50% smaller than the Doppler's. This could be linked to the fact that *Typhoon* measures apparent displacements, which are later converted to velocities (Sec. 2.c). Small-scales velocity structures, either in time or space, are less accurately perceived. Using a faster scan rate is likely to improve the results. Nevertheless, *Typhoon* performs better than the cross-correlation technique: the optimized algorithm presented in Hamada et al. (2015) recovers 39% of the TKE on the same dataset.

b. Spatial validation

During this phase of the experiment, the DL was colocated with the REAL. The REAL swept between 15° and 75° azimuth at 2° elevation every 17 s. The DL held its beam fixed at 45° azimuth and 2° elevation, measuring the radial velocity component as a function of range and time. DL measurements were integrated over one second, with a range gate of 48 m. The temporal resolution

---

[6] When dealing with circular data such as angles, the slope for the linear regression should be fixed to 1. The offset and R^2 only are computed, see e.g. Fisher (1995).

of DL measurements is therefore much finer than that of the REAL flow fields, and conversely for the spatial resolution (see Table 1).

Instead of holding the DL beam fixed, a PPI sweeping strategy identical to the REAL's could have been used, thus allowing the comparison of radial components over the whole scan domain. However, two arguments support the choice of a fixed beam:

- With a moving beam set-up, the integration time for DL measurements was reduced to less than 0.1 s. This would cause the SNR to decrease very rapidly. Typically in Chico the maximum range with useful data would be on the order of 1500 m, significantly below that of the REAL's.

- The radial velocity fields collected by the DL would suffer from the same distortions as the backscatter data (Section 4.c), so correcting these distortions would be challenging.

The data used for the spatial validation were recorded in December 2013 and January 2014. In Chico, CA, the days are shorter and the air is cleaner during this season than in the autumn when time-series data were collected. Both the DL and the REAL are affected. Data are of lower quality than shown for the temporal validation. The availability of 10-min averages falls below 50% after 3 km for both instruments and at 5 km it is below 5%. Therefore, the analysis is restricted to the first 3 km. Furthermore, it should be noted that the prevailing wind direction during this time over Chico, CA is northwesterly. At 45° azimuth, the line-of-sight component corresponds mostly to the *cross-stream*, turbulent wind perturbations. In these data, its magnitude remains mostly below 3 m s^{-1}. Figure 12 shows a comparison of radial velocity measured by the DL and extracted from the 2-component fields obtained by *Typhoon* for a 8-hour period starting 8 January 2014 at 17:00 UTC.

In order to compute statistics, radial velocities were averaged. First, spatial resolution are matched by averaging *Typhoon* velocities in space according to DL range gates, then 10-min time-averages are computed at every range. A scatter plot of these 10-min averages is presented in Fig. 13, along with linear regression slopes, R^2 coefficient and distribution of differences. These values were obtained from 8-hour periods (17:00 to 01:00 UTC) for 8 days of December 2013 and January 2014. The R^2 coefficient (panel d) decreases with the range and this is expected as both instruments are affected by the gradual reduction in SNR. R^2 remains above 0.95 over the first 1.5 km, then slowly decreases to about 0.8 at 3 km. The overall R^2 is 0.928. While the relation between *Typhoon* and DL velocities remains linear, the slope (panel c) increases with the range, from about 0.95 at 0.5 km to 1.3 at 3 km. Velocities obtained by the cross-correlation method show a similar trend (Hamada et al. 2015). This leads to a theory that these discrepancies are due to a mismatch in the actual elevation angles of the beams during this phase of the experiment, especially considering the unbiased results of the temporal validation. At a lower elevation angle and therefore lower altitude, the DL would measure lower velocities.

c. Spectral analysis

In this section, temporal and spatial power spectra of the velocity components produced by *Typhoon* are presented, with the objective of characterizing the filtering effect of the algorithm – in particular, in the spatial domain. The velocity data analyzed were collected during the daytime and within the turbulent lower atmospheric boundary layer.[7] Therefore, an inertial subrange in the power spectra of the actual velocity field is expected.

[7]RHI scans collected every 15 min by the REAL during the 15 days included in the analysis show that the maximum convective boundary layer height, that typically occurred in the afternoon, ranged from 300–1200 m AGL.

The spectra are computed in natural coordinates to account for the anisotropy of atmospheric boundary layer turbulence. The west-east and south-north wind velocity components are projected, according to the mean wind direction, as streamwise (u_s) and cross-stream (v_n) components, such that u_s carries the mean speed and v_n has a null mean. The mean wind vector is defined accordingly to the investigated dimension, either in time or space. The spectra are finally averaged together according to the mean wind speed, using bins of 0–4 m s^{-1}, 4–8 m s^{-1}, 8–12 m s^{-1} and 12–16 m s^{-1}, in order to exibit their evolution with increasing wind speed and turbulent kinetic energy. The resulting power spectral densities (S) are multiplied by frequency (f) or wavenumber squared (κ^2) so that an inertial subrange would appear as a $-2/3$ slope and white noise would appear as $+1$ slope.

1) Temporal power spectra

During the experiment, the REAL collected PPI scans every 17 s and one RHI scan every 15 min. The RHI scan resulted in an 30 s interruption of the PPI scan sequence. The scan strategy of the DL provided vertical profiles of horizontal winds every 15$\pm$1 s. Since the FFT requires data points at a uniform time interval, the *Typhoon* and DL wind measurements were interpolated to a 5 s time series. From the 5 s time series data, we computed power spectra over consecutive 10 min intervals. The 10-min mean wind vector was used for the projection in natural coordinates and the binning of spectra, as defined above. The resulting spectra have a Nyquist frequency of 0.029 Hz (34 s period) for the *Typhoon* velocities and 0.033 Hz (30 s period) for the DL. The lowest frequency is 1.67×10^{-3} Hz (10 min period).

The spectra are presented in Fig. 14. Those from the Doppler lidar are consistently higher than the spectra from *Typhoon*, this is consistent with our observation that the TKE measured from Doppler velocities are larger than those from *Typhoon* (Fig. 10). The temporal spectra appear to become flatter as the mean wind speed increases. We hypothesize that this may be caused by the challenges that both *Typhoon* and the DL face under windy conditions. For the DL, increased variability of the actual wind velocity field in the VAD sample area results in more error in the horizontal wind vector estimate. The increased error appear as noise at these time scales and flatten the spectrum. For *Typhoon*, windy conditions result in larger horizontal displacements between scans and faster deformation of aerosol coherent structures.

2) Spatial power spectra

An independent observation of the 2-component 2-D velocity field does not exist for comparison with those produced by *Typhoon*. A dual-Doppler lidar set up could have provided it, but would have doubled the cost and complexity of the project. Therefore, to investigate the integrity of the vector flow fields in space, spatial power spectra are considered.

A 1 km diameter circular area is considered, centered on the DL at 1.53 km range. All of the vectors within this area (from a single flow field in time) are used to compute the spatial mean wind vector, which then define a natural coordinate system. Vectors of the flow field are interpolated on a 128×128 point grid (1024 m $\times$ 1024 m) that is centered on the DL and aligned with the natural coordinate system, and then projected as streamwise (u_s) and cross-stream (v_n) components. This operation was performed for each flow field independently and resulted in 30092 flows fields over 15 days. At 4° elevation, the 1024 m $\times$ 1024 m area covers a range of altitudes from about 50 m to 150 m AGL. A possible impact of this is that the turbulence statistics within this sloped domain are slightly inhomogeneous. Nevertheless, for each component u_s, v_n, the 2D power spectral densities computed by FFT for each flow field are averaged together according to the mean spatial wind speed. Finally, slices of the resulting 2D power spectra were extracted along the streamwise and cross-stream directions. This results in four 1D spectra for each wind speed bin: along the streamwise and cross-stream directions, for each of the streamwise and cross-stream components. The Nyquist wavenumber is $\kappa/2\pi = 0.0625$ m^{-1} (16 m wavelength), the lowest wavenumber is 9.77×10^{-4} m^{-1} (1204 m).

The spectra in the top row of Fig. 15 show the TKE increasing as expected as function of wind speed. Each spectrum has a maximum amplitude at low wavenumbers. We hypothesize that the peak corresponds to one over the Eularian length scale, and is within the energy containing range (Kaimal and Finnigan 1994). However, the spectra are steeper than $\kappa^{-2/3}$. We attribute this to two factors. First is the likely absence of aerosol features at all scales and all locations in the scan area at all times. Second is the regularization used in *Typhoon* which favors a smooth motion field, especially as the estimation reaches the smallest scales.

A transfer function describes the ratio of two spectra and, in the present work, represents the attenuation of the actual wind field caused by the motion estimation as a function of wavenumber. A highly idealized spectrum is constructed to serve as the reference. This is done by first locating the maximum of each mean spatial spectra shown in Fig. 15. We assume that the observed power at wavenumbers smaller than the peak in the spectra are accurately captured by the algorithm and serve as a proper approximation of the power at those large scales. For scales smaller than the peak, we extrapolate by a power-law dependence through the higher wavenumbers that mimics the inertial subrange (a $\kappa^{-2/3}$ spectrum). The transfer functions are then given by the ratio of the observed mean spectra over the idealized spectrum, and presented in the bottom row of Fig. 15. The higher the wind speed, the more energy is missing at small scales. The ratio typically drops below 50% at scales of $\approx$100 m ($\kappa/2\pi \approx 0.01$ m^{-1}) for the highest wind speeds, and $\approx$75 m for the lowest.

6. Broader perspectives and conclusions

In a recent paper, entitled *Review of turbulence measurements using ground-based wind lidars*, Sathe and Mann (2013) conclude that "Non-coherent detection may also provide possible new ways to estimate atmospheric turbulence, but to our knowledge it does not, so far, challenge the capabilities of coherent Doppler lidars." In this paper, we have (1) introduced a new motion estimation method; (2) made the first direct comparisons of the "non-Doppler motion estimation approach" with Doppler lidar; and (3) computed transfer functions to estimate the filtering effect of the approach. The new motion estimation method resolves finer spatial scale flow details than the traditional cross-correlation algorithm (Hamada et al. 2015). The comparisons in the time domain reveal excellent correlation in terms of 10-min averages, close for example to standards expected of commercial floating lidars (Carbon Trust 2013). However, the proposed approach still underestimates the TKE by about 50% of what is observed by Doppler lidar. It is important to keep in mind that the Doppler also provides a filtered version of the actual flow field.

Two horizontal components are required for wind speed and direction. The proposed approach delivers dense 2-component wind fields from a single lidar, whereas a single Doppler only produces a single component. In addition to wind resource assessment, wind fields such as delivered by *Typhoon* from REAL imagery enable the visualization and investigation of meteorological phenomena such as vortices and fronts. They also open the possibility of studies in the Lagrangian reference frame, and the tracking of flow structures or aerosol features.

Acknowledgments. This material is based upon work supported by the National Science Foundation's Physical and Dynamic Meteorology Program under Grant No. AGS 1228464. Dr. Christoph Thomas provided helpful suggestions for data analysis.

APPENDIX A

Mathematical Symbols

- $\forall$ for all;
- $\subset$ subset of;
- $\in$ in (belonging to);
- $\mathbb{N}$, $\mathbb{R}$ the sets of natural and real numbers, respectively.

APPENDIX B

Parameters of *Typhoon*

Unless specified, results were obtained using the following parameters for *Typhoon*:

- version: cuTyphoon 1.0;
- wavelet basis: Daubechies, 10 vanishing moments;
- wavelet scales: 8 details scales considered and estimated;
- pyramid steps=1, scaling factor=50%;
- data model: DFD, smoothing kernel $\sigma = 0.5$;
- regularization: Horn & Schunk, $\alpha = 0.05$;
- data range: $[-0.5, 0.5]$, with normalization, without histogram matching.

References

Adrian, R. J., 2005: Twenty years of particle image velocimetry. *Exp. Fluids*, **39 (2)**, 159–169.

Adrian, R. J., and J. Westerweel, 2010: *Particle image velocimetry*, Vol. 30. Cambridge University Press.

Axel, A., and J. Ailt-Wiard, 2014: Calibration certificate – HALO Photonics Stream Line 1012-55. Tech. rep., Deutsche WindGuard.

Bailey, B. H., Ed., 2012: *Wind resource assessment: a practical guide to developing a wind project*. John Wiley & Sons.

Carbon Trust, 2013: Roadmap for the commercial acceptance of floating LIDAR technology. Tech. Rep. CTC819, Offshore Wind Accelerator programme, 30 pp.

Corpetti, T., E. Mémin, and P. Pérez, 2002: Dense estimation of fluid flows. *IEEE Trans. Pattern Anal. Mach. Intell.*, **24 (3)**, 365–380.

Dérian, P., 2012: Wavelets and Fluid Motion Estimation. Ph.D. thesis, MATISSE, Université Rennes 1, 100 pp., [Available online at http://tel.archives-ouvertes.fr/tel-00761919/PDF/theseDERIAN_v3_BU_.pdf.].

Dérian, P., P. Héas, C. Herzet, and E. Mémin, 2013: Wavelets and optical flow motion estimation. *Num. Math. Theor. Meth. Appl.*, **6 (1)**, 116–137.

Dérian, P., P. Héas, E. Mémin, and S. Mayor, 2010: Dense motion estimation from eye-safe aerosol lidar data. *International Laser Radar Conference (ILRC25)*, St Petersburg, Russia, Vol. 1, 377–380, oral presentation S3O-04.

Dugan, J., S. Anderson, C. Piotrowski, and S. Zuckerman, 2014: Airborne infrared remote sensing of riverine currents. *IEEE Trans. Geosci. Rem. Sens.*, **52 (7)**, 3895–3907, doi:10.1109/TGRS.2013.2277815.

Eloranta, E., J. King, and J. Weinman, 1975: The determination of wind speeds in the boundary layer by monostatic lidar. *J. Appl. Meteor.*, **14 (8)**, 1485–1489.

Fisher, N. I., 1995: *Statistical analysis of circular data*. Cambridge University Press.

García-Pereda, J., and R. Borde, 2014: The impact of the tracer size and the temporal gap between images in the extraction of atmospheric motion vectors. *J. Atmos. Ocean. Technol.*, **31 (8)**, 1761 – 1770, URL http://mantis.csuchico.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=97381913&site=eds-live.

Hamada, M., 2014: Evaluations of the performance of a cross-correlation algorithm for wind velocity estimation using syntetic backscatter lidar images and velocity fields. M.S. thesis, Environmental Sciences, California State University, Chico, 159 pp., [Available online at http://phys.csuchico.edu/lidar/publications/Masaki_Hamada_MS_thesis.pdf.].

Hamada, M., P. Dérian, C. F. Mauzey, and S. D. Mayor, 2015: Optimization of the cross-correlation algorithm for two-component wind field estimation from single aerosol lidar data and comparison with Doppler lidar. *Submited to J. Atmos. Ocean. Technol.*

Held, A., T. Seith, I. M. Brooks, S. J. Norris, and S. D. Mayor, 2012: Intercomparison of lidar aerosol backscatter and in-situ size distribution measurements. *European Aerosol Conference*, Granada, Spain, presentation number B-WG01S2P05.

Horn, B. K., and B. G. Schunck, 1981: Determining optical flow. *Artif. Intell.*, **17**, 185–203.

Kadri Harouna, S., P. Dérian, P. Héas, and E. Mémin, 2013: Divergence-free Wavelets and High Order Regularization. *Int. J. Computer Vision*, **103**, 80–99, URL http://hal.archives-ouvertes.fr/hal-00646104.

Kaimal, J. C., and J. J. Finnigan, 1994: *Atmospheric Boundary Layer Flows*. Oxford University Press, 200 Madison Ave., New York, NY 10016, 289 pp.

Kanak, K. M., 2005: Numerical simulation of dust devil-scale vortices. *Quart. J. Roy. Meteor. Soc.*, **131 (607)**, 1271–1292.

Leese, J., C. Novack, and B. Clark, 1971: An automated technique for obtained cloud motion from geosynchronous satellite data using cross correlation. *J. Appl. Meteor.*, **10**, 118–132.

Mallat, S., 2008: *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press.

Mann, J., A. Pena, F. Bingol, R. Wagner, and M. S. Courtney, 2010: Lidar scanning of momentum flux in and above the atmospheric surface layer. *J. Atmos. Ocean. Technol.*, **27**, 959–976.

Mann, J., and Coauthors, 2009: Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer. *Meteor. Z.*, **18**, 135–140.

Mauzey, C. F., P. Dérian, and S. D. Mayor, 2014: Wavelet-based optical flow for real-time wind estimation using CUDA. *GPU Technology Conference*, San Diego, CA, poster P4253.

Mauzey, C. F., J. P. Lowe, and S. D. Mayor, 2012: Real-time wind velocity estimation from aerosol lidar data using graphics hardware. *GPU Technology Conference*, San Diego, CA, poster AV10.

Mayor, S. D., and E. W. Eloranta, 2001: Two-dimensional vector wind fields from volume imaging lidar data. *J. Appl. Meteor.*, **40 (8)**, 1331–1346.

Mayor, S. D., J. P. Lowe, and C. F. Mauzey, 2012: Two-component horizontal aerosol motion vectors in the atmospheric surface layer from a cross-correlation algorithm applied to scanning elastic backscatter lidar data. *J. Atmos. Ocean. Technol.*, **29 (11)**, 1585–1602.

Mayor, S. D., and S. M. Spuler, 2004: Raman-shifted eye-safe aerosol lidar. *Appl. Optics*, **43 (19)**, 3915–3924.

Mayor, S. D., S. M. Spuler, B. M. Morley, and E. Loew, 2007: Polarization lidar at 1.54 μm and observations of plumes from aerosol generators. *Opt. Eng.*, **46 (9)**, 096 201–096 201.

Melville, W. K., and P. Matusov, 2002: Distribution of breaking waves at the ocean surface. *Nature*, **417 (6884)**, 58–63.

Päschke, E., R. Leinweber, and V. Lehmann, 2014: A one year comparison of 482 mhz radar wind profiler, rs92-sgp radiosonde and 1.5 μm Doppler lidar wind measurements. *Atmos. Meas. Tech. Discuss.*, **7 (11)**, 11 439–11 479.

Patton, E. G., and Coauthors, 2011: The canopy horizontal array turbulence study. *Bull. Amer. Meteor. Soc.*, **92 (5)**, 593–611.

Pearson, G., F. Davies, and C. Collier, 2009: An analysis of the performance of the UFAM pulsed Doppler lidar for observing the boundary layer. *J. Atmos. Ocean. Technol.*, **26**, 240–250.

Sasano, Y., H. Hirohara, T. Yamasaki, H. Shimizu, N. Takeuchi, and T. Kawamura, 1982: Horizontal wind vector determination from the displacement of aerosol distribution patterns observed by a scanning lidar. *J. Appl. Meteor.*, **21 (10)**, 1516–1523.

Sathe, A., and J. Mann, 2012: Measurement of turbulence spectra using scanning pulsed wind lidars. *J. Geophys. Res.*, **117 (D1)**, doi:10.1029/2011JD016786.

Sathe, A., and J. Mann, 2013: A review of turbulence measurements using ground-based wind lidars. *Atmos. Meas. Tech. Discuss.*, **6 (4)**, 6815–6871.

Sathe, A., J. Mann, J. Gottschall, and M. S. Courtney, 2011: Can wind lidars measure turbulence? *J. Atmos. Ocean. Technol.*, **28**, 853–868.

Scambos, T. A., M. J. Dutkiewicz, J. C. Wilson, and R. A. Bindschadler, 1992: Application of image cross-correlation to the measurement of glacier velocity using satellite image data. *Remote Sens. Environ.*, **42 (3)**, 177–186.

Schmetz, J., K. Holmund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, and L. V. D. Berg, 1993: Operational cloud-motion winds from Meteosat infrared images. *J. Appl. Meteor.*, **32 (7)**, 1206–1225.

Schmidt, H., and U. Schumann, 1989: Coherent structure of the convective boundary layer derived from large-eddy simulations. *J. Fluid Mech.*, **200**, 511–562.

Schols, J., and E. Eloranta, 1992: Calculation of area-averaged vertical profiles of the horizontal wind velocity from volume-imaging lidar data. *J. Geophys. Res.*, **97 (D17)**, 18 395–18 407.

Spuler, S. M., and S. D. Mayor, 2005: Scanning eye-safe elastic backscatter lidar at 1.54 μ m. *J. Atmos. Ocean. Technol.*, **22 (6)**, 696–703.

Stawiarski, C., K. Träumner, C. Knigge, and R. Calhoun, 2013: Scopes and challenges of dual-Doppler lidar wind measurements—an error analysis. *J. Atmos. Ocean. Technol.*, **30 (9)**, 2044–2062.

Stumpf, A., J. Malet, P. Allemand, G. Skupinski, and M. Pierrot-Deseilligny, 2013: Terrestrial multi-view photogrammetry for landslide monitoring. *AGU Fall Meeting Abstracts*, A5.

Sullivan, P. P., and E. G. Patton, 2011: The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulation. *J. Atmos. Sci.*, **68 (10)**, 2395–2415.

Van Dyke, M., 1982: *An album of fluid motion*. Parabolic Press Stanford, CA.

Weber, B., D. Wuertz, D. Welsh, and R. McPeek, 1993: Quality controls for profiler measurements of winds and rass temperatures. *J. Atmos. Ocean. Technol.*, **10 (4)**, 452–464.

LIST OF TABLES

TABLE 1. Main parameters of DL and REAL measurements for the temporal and spatial validation experiments.

	Temporal validation		Spatial validation	
	Doppler	REAL	Doppler	REAL
scan type	VAD	PPI	STARE	PPI
azimuth (°)	–	$[-15;45]$	45	$[15;75]$
elevation (°)	–	4	2	2
range (km)	–	$[0.5;5.5]$	$[0;5]$	$[0.5;5.5]$
components	2	2	1	2
δx (m)	–	8	48	8
δt (s)	15 ± 1	17	1	17

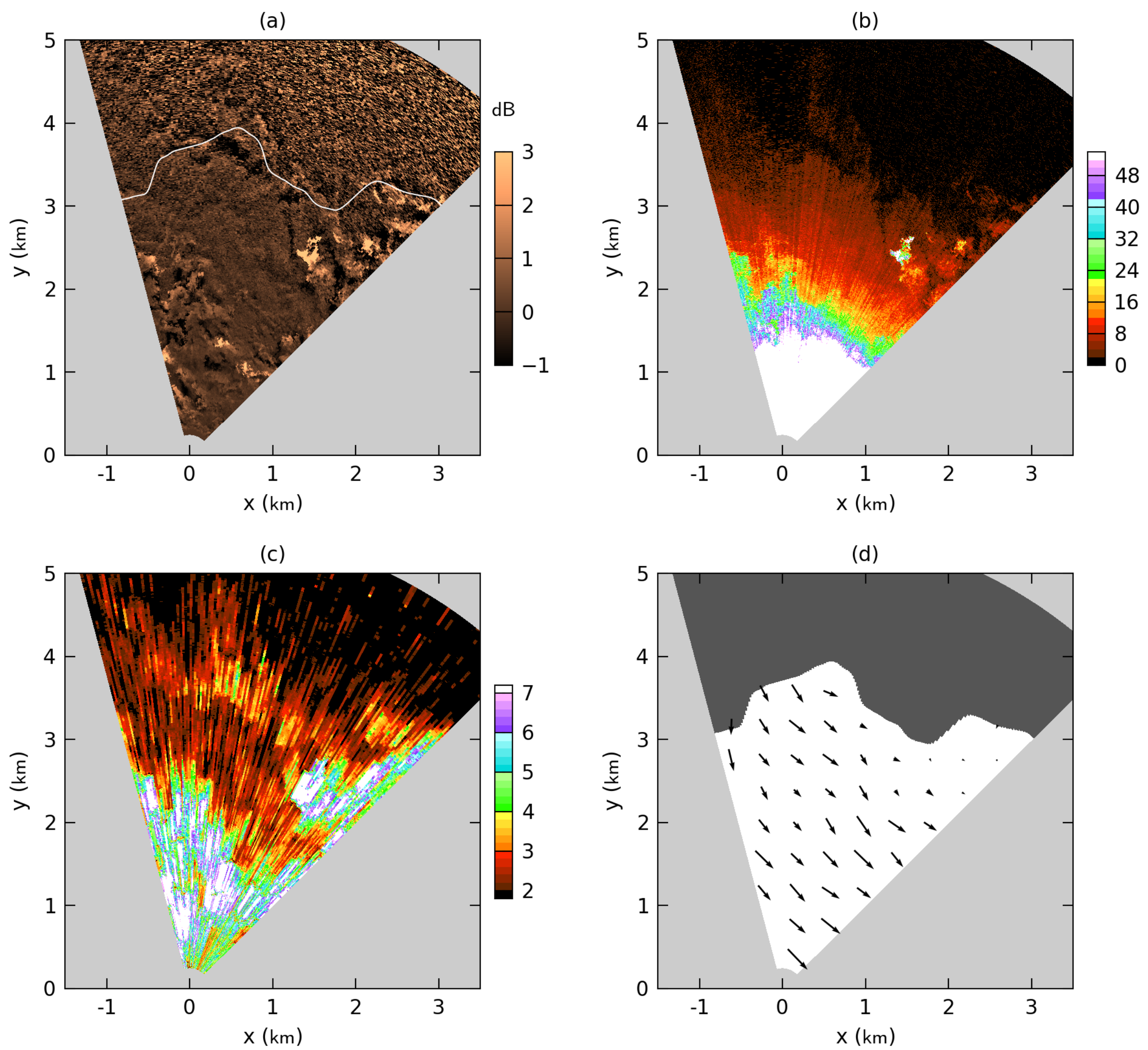
TABLE 2. RMS error, linear regression variables (slope, offset), correlation coefficient R^2, number of points and recovery percentage w.r.t. DL reference for the 10-min averaged wind component u (west-east), for the temporal validation results (Sec. 5.a).

case	RMSE (m s^{-1})	slope	offset (m s^{-1})	R^2	# points	% recovery
light	0.17	1.047	-0.01	0.951	61	84.7
moderate	0.29	0.974	-0.05	0.979	72	100
strong	0.33	0.938	0.32	0.984	72	100
15 days	0.29	0.989	-0.03	0.995	892	99.1

TABLE 3. RMS error, linear regression variables (slope, offset), correlation coefficient R^2, number of points and recovery percentage w.r.t. DL reference for the 10-min averaged wind component v (south-north), for the temporal validation results (Sec. 5.a).

case	RMSE (m s^{-1})	slope	offset (m s^{-1})	R^2	# points	% recovery
light	0.27	0.660	-0.02	0.600	61	84.7
moderate	0.23	0.999	0.00	0.991	72	100
strong	0.34	0.897	-0.72	0.929	72	100
15 days	0.29	1.001	0.03	0.997	892	99.1

LIST OF FIGURES



835 FIG. 1. Example of preprocessing applied to a horizontal scan collected on 3 October 2013 at 23:14:10 UTC. Panel (a) is the preprocessed backscatter data. Panel (b) is the raw SNR (7), revealing a $1/r^2$ decay. Panel (c) is the image SNR (8) computed using a 384 m window. Panel (d) is the valid data domain computed from image SNR. Motion is estimated in the white area only, excluding far-range noisy regions. The far-range boundary (9) of this area is also shown in (a) as a white line. Resulting, decimated vector flow field has been added to the valid area in (d).

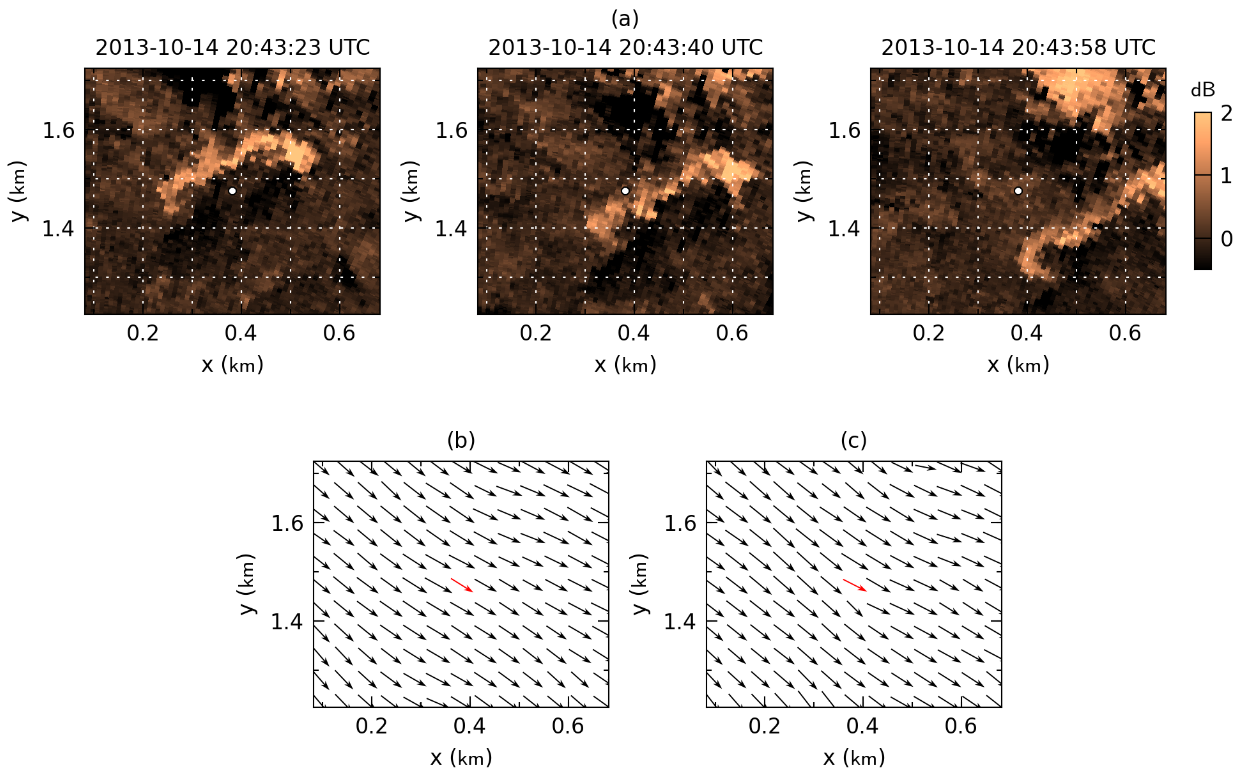


FIG. 2. Illustration of the experimental design for the temporal validation of motion estimation vectors. Panel (a) is a short sequence of 3 consecutive PPI scans collected on 14 October 2013 by the REAL. The displayed area is a close-up centered on the Doppler lidar (white marker) used for validation. The copper shading indicates the intensity, in dB, of aerosol backscatter. A large aerosol feature is being advected south-east and passes over the DL. Panel (b), and (c) show the velocity fields estimated by *Typhoon* (black arrows) from each pair of scans; they were decimated by a factor of 6 along both dimensions for the sake of visualization. Measurements from the DL (red arrows) at 100 m AGL show a good agreement with estimates, with a wind speed of ≈ 5.4 m s^{-1}.

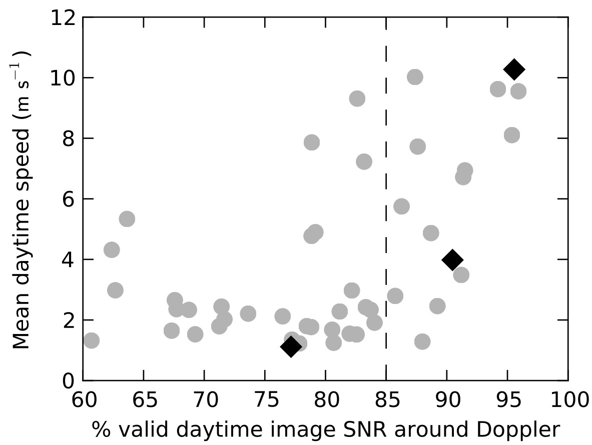
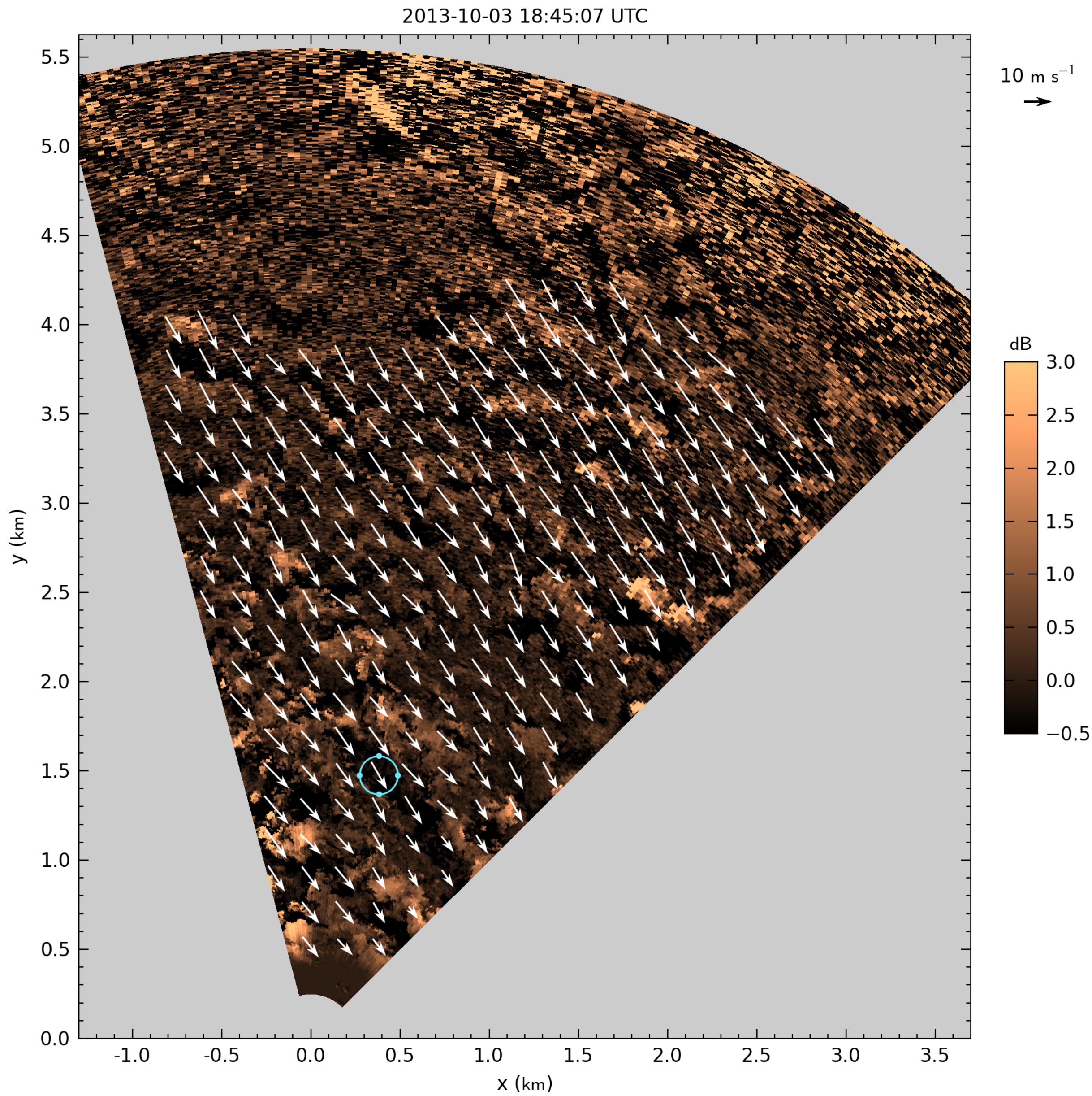
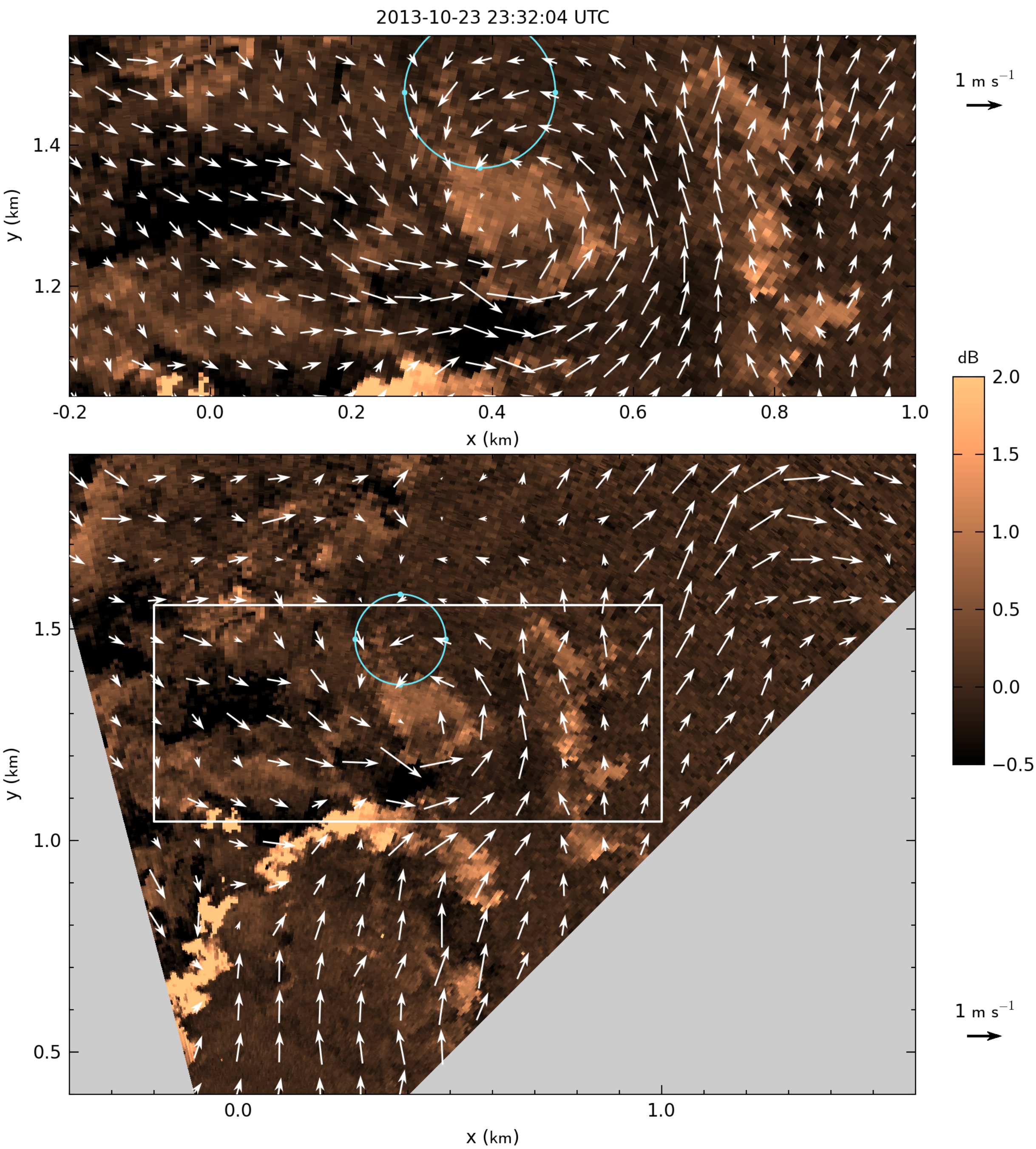


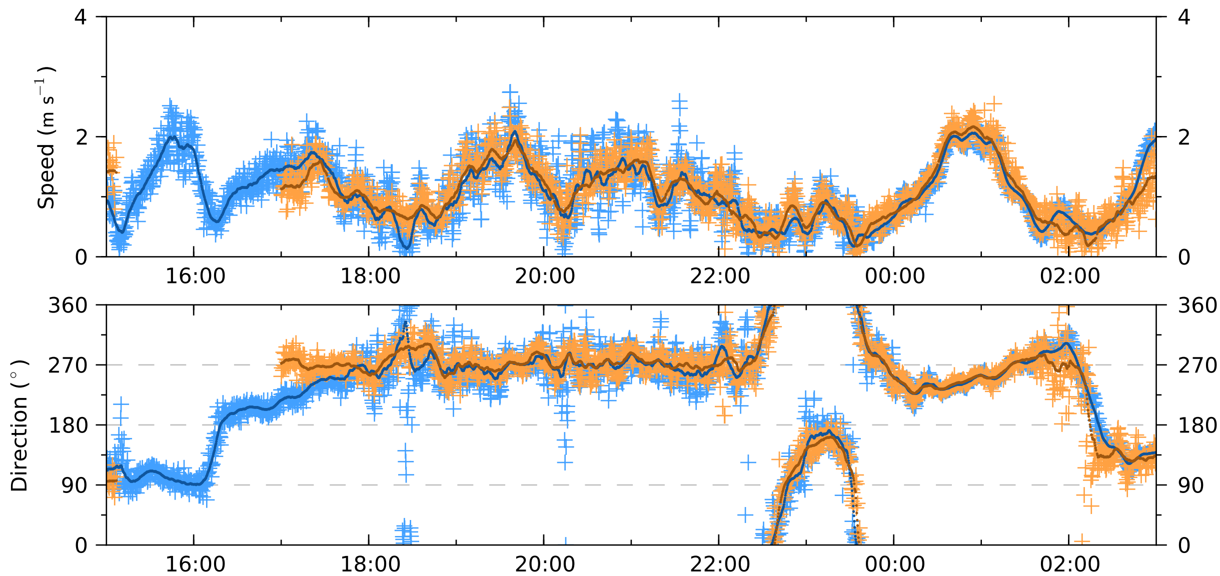
FIG. 3. Distribution of days in terms of valid image SNR (Sec. 4.b) in a 50 m radius around DL location (horizontal axis) versus mean wind speed measured by the DL at 100 m AGL (vertical axis), during daytime. Days for which time-series are presented (Fig. 6, 7, 8) are represented with a black diamond. A total of 57 days are considered, of which 9 have less than 60% valid SNR_{img} and are not visible here. The 15 days having more than 85% valid SNR_{img} were investigated for the statistics shown in Fig. 9.



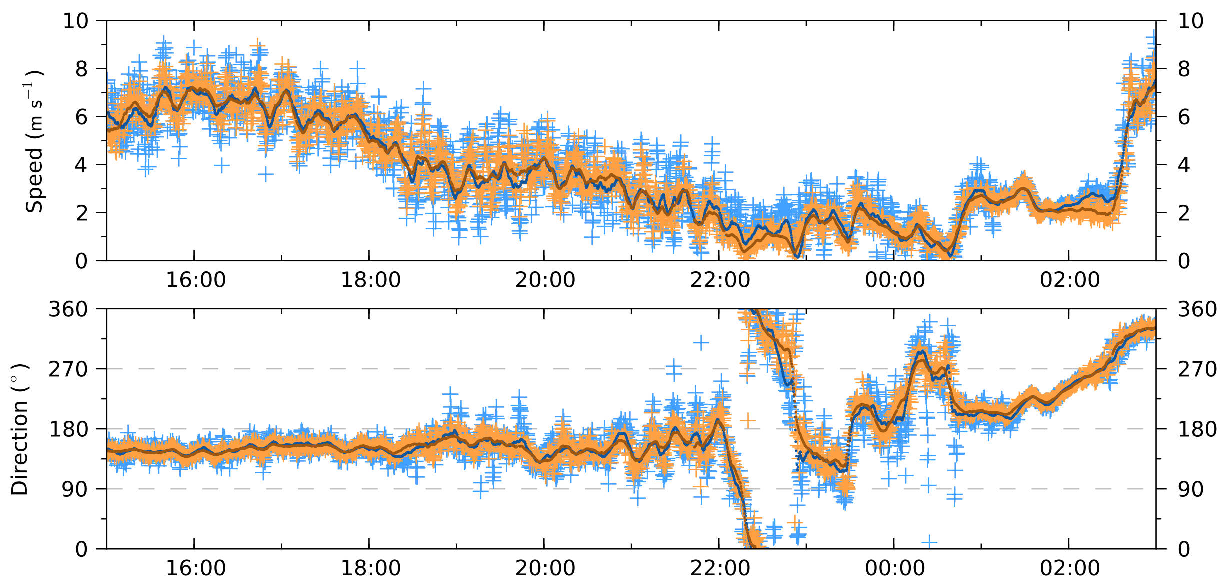
853 FIG. 4. Wind field obtained by *Typhoon* 3 October 2013 at 18:45:07 UTC, superimposed on the first scan
854 of the pair used for estimation. Wind velocity was ≈ 14 m s^{-1}. The motion field was decimated along both
855 dimensions by a factor of 24. The turquoise circle represents the cone section sampled by the DL during the
856 VAD scan.



857 FIG. 5. Wind field obtained by *Typhoon* 23 October 2013 at 23:32:04 UTC, superimposed on the first scan of
858 the pair used for estimation. The upper panel shows a close up on a vortex of radius $\approx$ 200 m. The motion field
859 was decimated along both dimensions by a factor of 6 and 12 for the top and bottom panels, respectively. The
860 turquoise circle represents the cone section sampled by the DL during the VAD scan.



861 FIG. 6. Times series of wind speed (top) and direction (bottom) as measured by the DL (blue) and estimated by proposed method (orange), for a 12-hour period starting 23 October 2013 at 15:00 UTC (light wind case). Light + markers are instantaneous values, darker lines are the 10-min rolling averages. The rapid change in direction is the signature of the vortex presented in Fig. 5.



865 FIG. 7. Times series of wind speed (top) and direction (bottom) as measured by the DL (blue) and estimated
866 by proposed method (orange), for a 12-hour period starting 17 September 2013 at 15:00 UTC (moderate wind
867 case). Light + markers are instantaneous values, darker lines are the 10-min rolling averages.

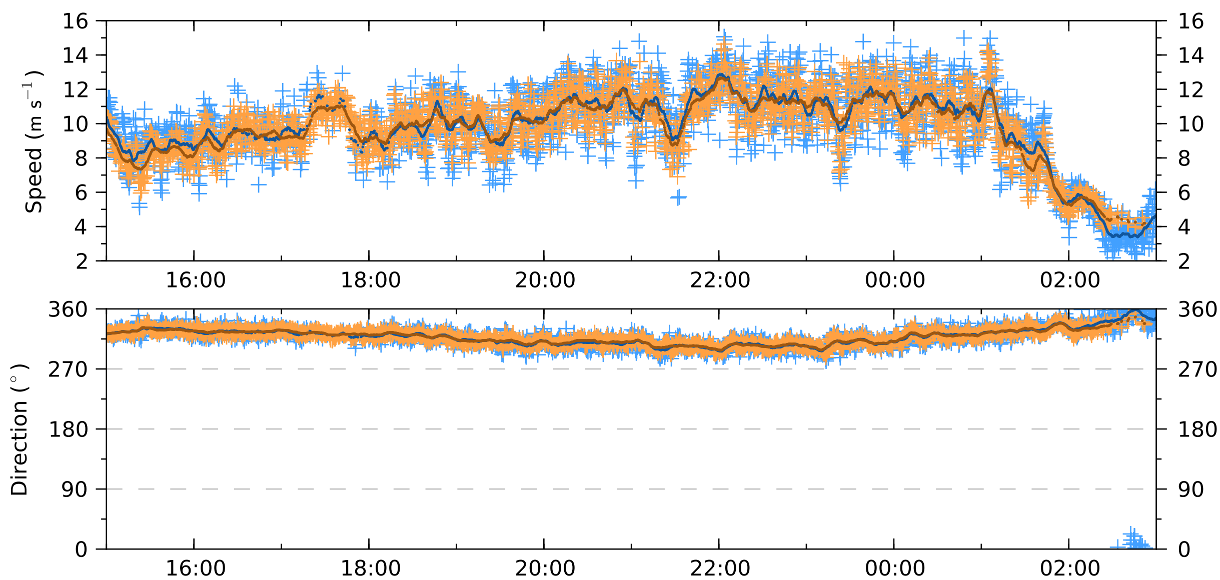


FIG. 8. Times series of wind speed (top) and direction (bottom) as measured by the DL (blue) and estimated by proposed method (orange), for a 12-hour period starting 9 October 2013 at 15:00 UTC (strong wind case). Light + markers are instantaneous values, darker lines are the 10-min rolling averages.

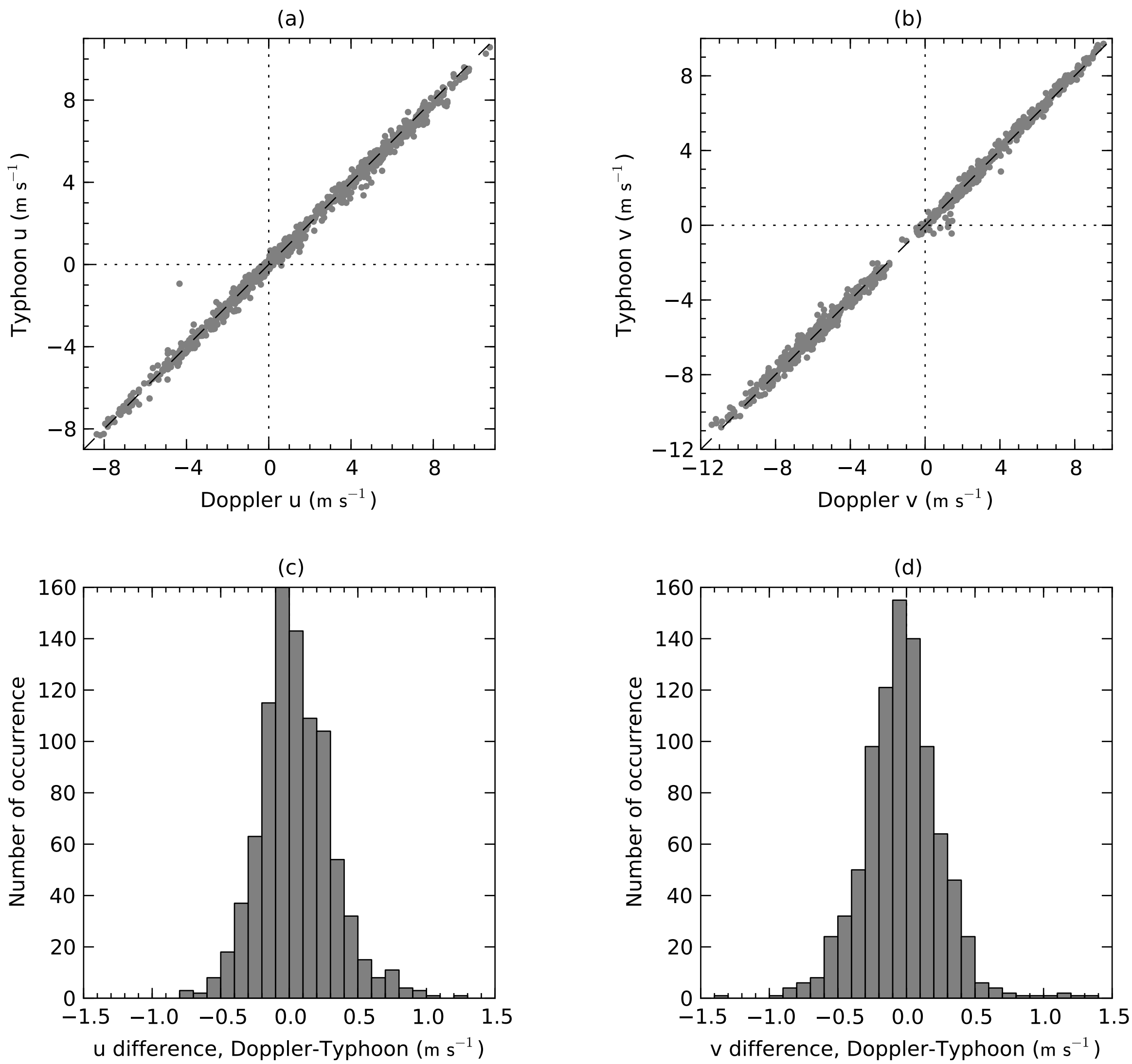


FIG. 9. Panels (a) and (b) are scatter plots of 10-min averaged wind components *u* and *v* measured by the DL at 100 m AGL (horizontal axis) versus estimated by *Typhoon* (vertical axis), combining the 15 days having $> 85\%$ valid SNR_{img} during daytime (Fig. 3) – 892 points total. Panels (c) and (d) are the distribution of differences for the same dataset.

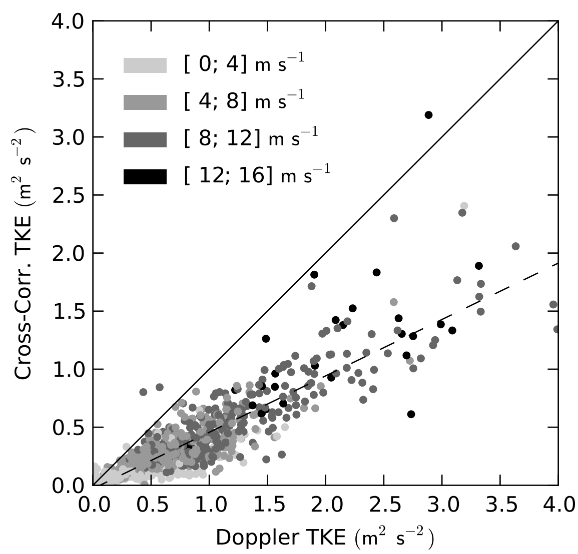


FIG. 10. Scatter plot of the TKE measured over 10-min intervals, by the DL at 100 m AGL (horizontal axis) versus estimated by the proposed method (vertical axis) – 892 points total. The gray shading indicated the mean wind speed measured over the interval. A linear regression (dashed line) gives a slope of 0.49 and offset of -0.03.

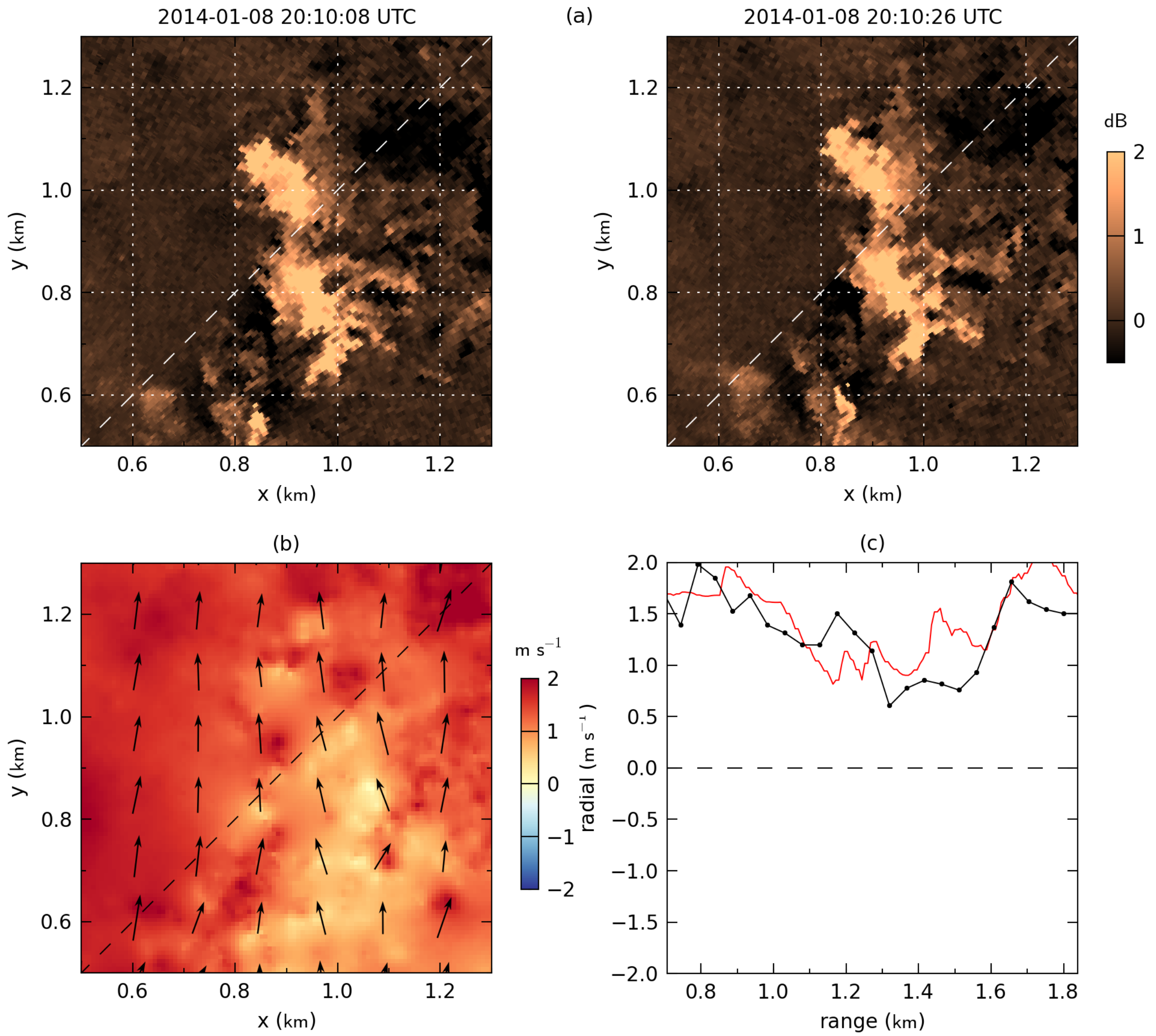
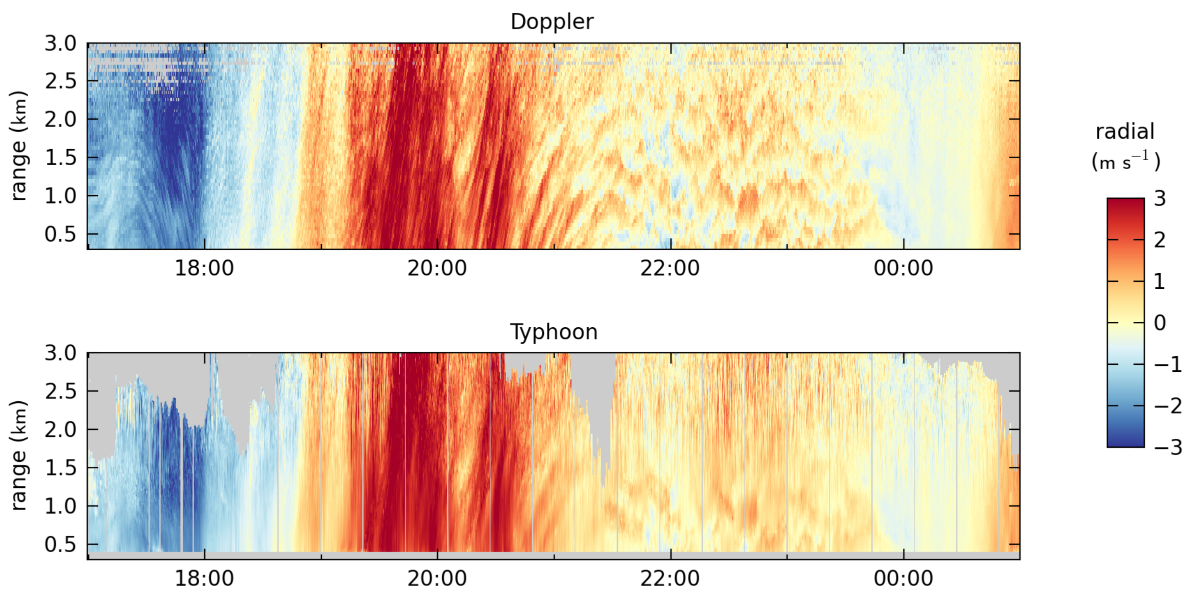


FIG. 11. Illustration of the experimental design for the spatial validation of motion estimation vectors. Panel (a) shows subsets of 2 consecutive PPI scans collected on 8 January 2014 by the REAL. The displayed area is a close-up centered on the DL line-of-sight at 45° azimuth (dashed white line). The copper shading indicates the intensity, in dB, of aerosol backscatter. A large aerosol feature is being advected north. Panel (b) shows the velocity field (black arrows) estimated by *Typhoon* from these two scans; the vector field was decimated by a factor of 15 along both dimensions for the sake of visualization. The color shading in the background indicates the corresponding radial velocity. Panel (c) compares the radial velocities measured by the Doppler (black line) and extracted from the 2-component field estimated by *Typhoon* (red line), along the DL line-of-sight.



886 FIG. 12. Comparison of *radial wind component* at 45° azimuth and 2° elevation measured by the DL (top) and estimated by proposed method (bottom), as a function of time (horizontal axis) and range (vertical axis), for a 8-hour period starting 8 January 2014 at 17:00 UTC. Gray shading indicates missing or discarded data.

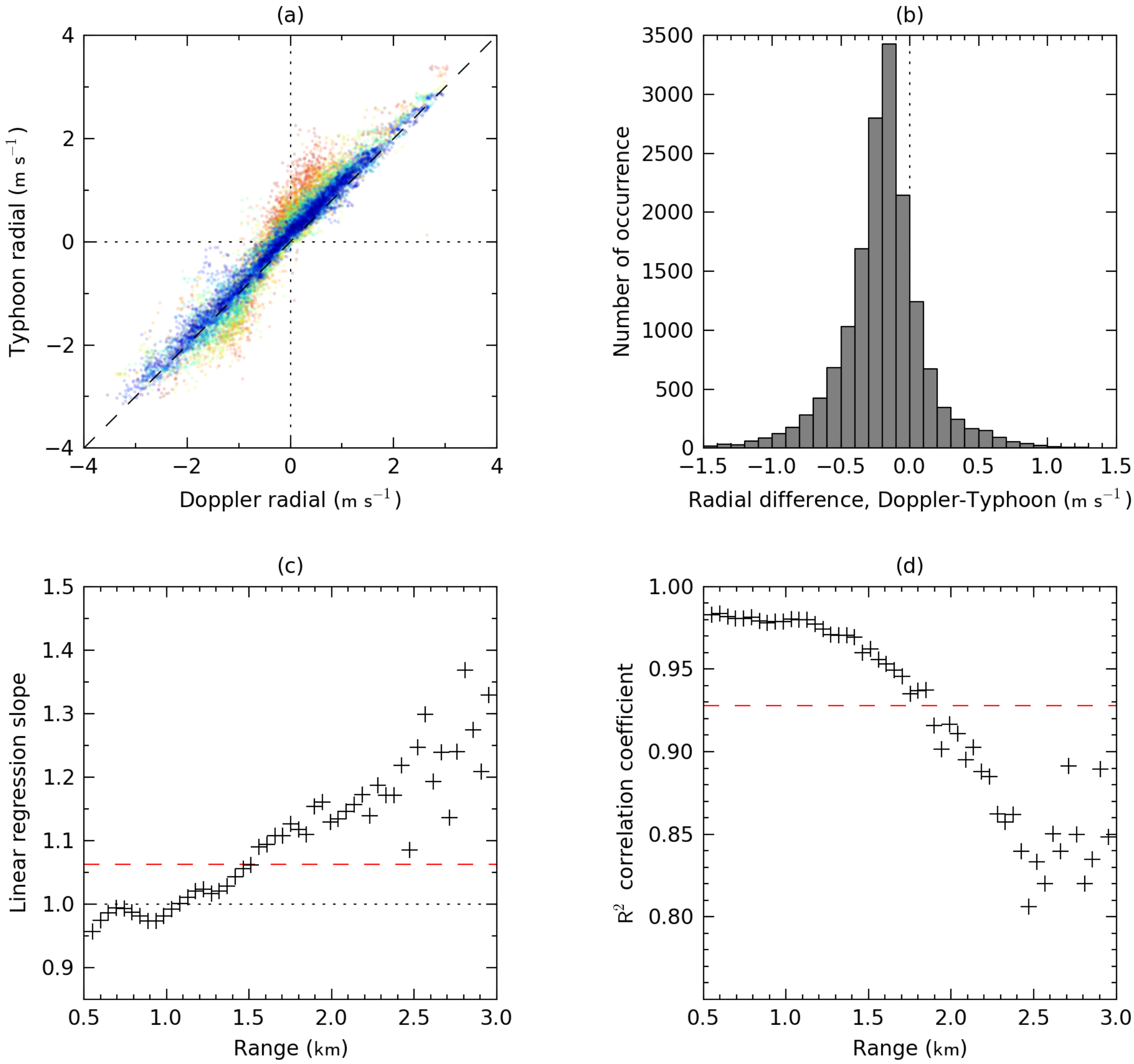


FIG. 13. Panel (a), scatter plot of 10-min averaged radial wind component measured by the DL (horizontal axis) versus estimated by the proposed method (vertical axis). Color indicates the range, from blue (0.5 km) to red (3 km). Panel (b), histogram of differences. Panel (c), slope of linear regression as a function of range. Panel (d), R^2 coefficient as a function of range. Dashed red lines indicate overall slope and R^2 values.

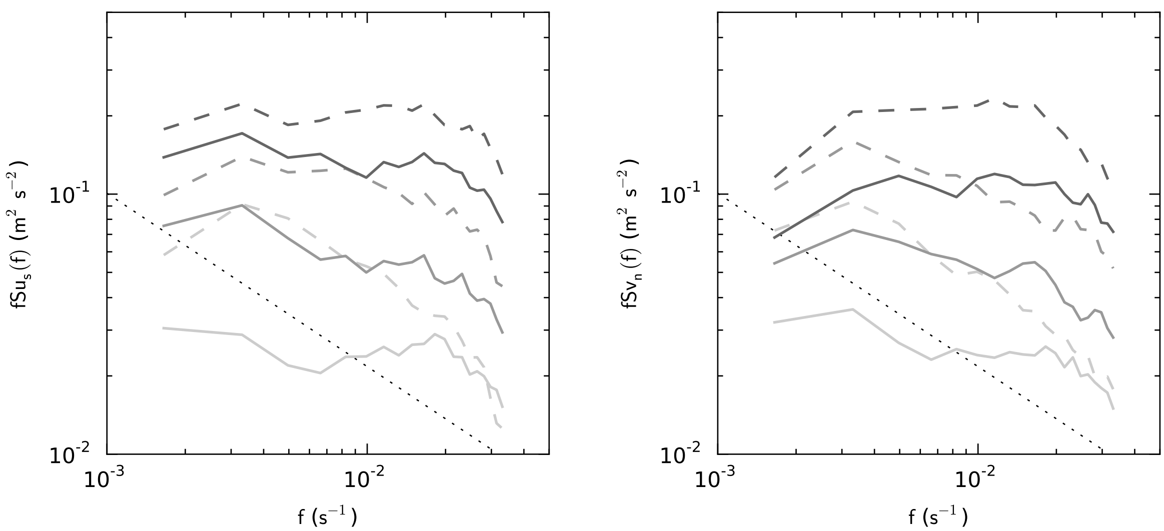


FIG. 14. Temporal spectra for stream-wise component u_s (left) and cross-stream component v_n (right) obtained by *Typhoon* (solid lines) and the DL (dashed lines). The shadings from light to dark gray correspond to wind speed ranges of [0;4], [4;8], and [8;12] m s^{-1}. The dotted line represents the -2/3 slope of the inertial subrange predicted by theory.

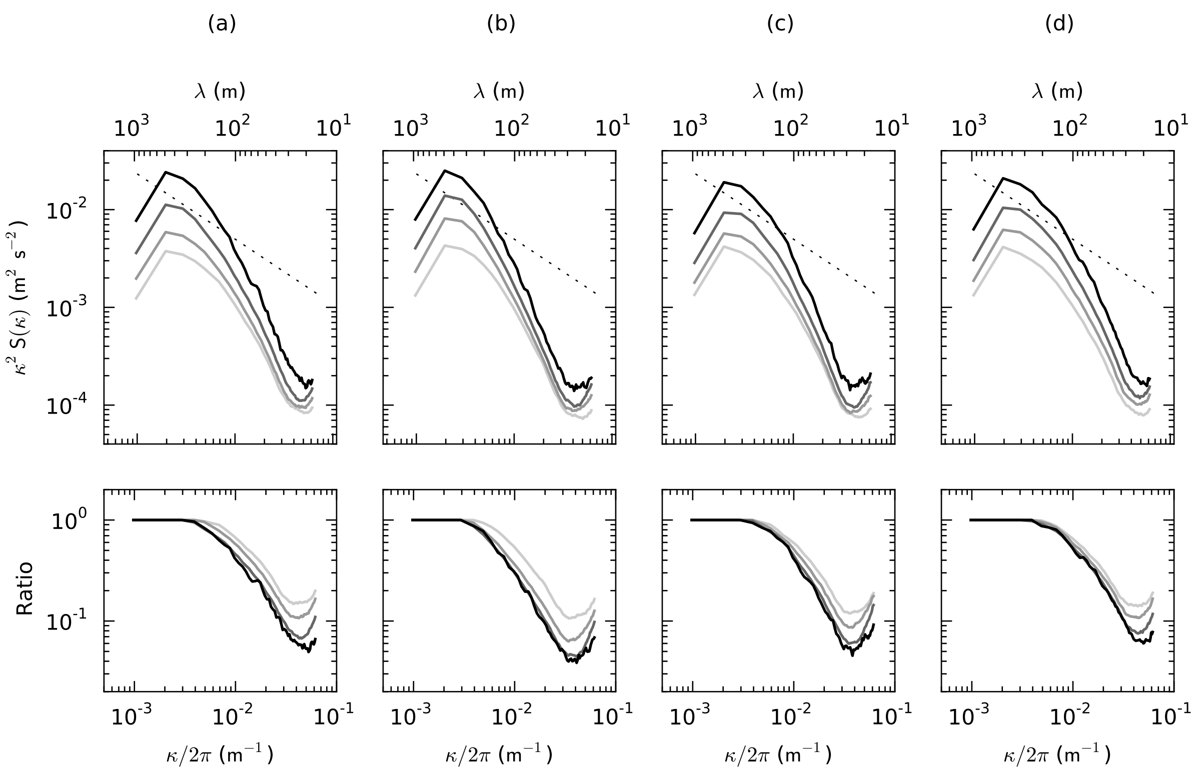


FIG. 15. Slices of 2D power spectral density (top) and corresponding transfer functions (bottom), for streamwise component u in the streamwise (a) and cross-stream (b) directions, and cross-wise component v in the streamwise (c) and cross-stream (d) directions. The shadings from light gray to black correspond to wind speed ranges of [0;4], [4;8], [8;12] and [12;16] m s^{-1}. The dotted line represents the -2/3 slope of the inertial subrange predicted by theory.