WindCube Data Filtering Guidelines

Revision History

Version	Status	Date/Name	Description of Change	Approver
0.1	Draft	23/11/2021 / Black Andrew	Initial Draft	--
0.2	Draft	30/11/2021 / Black Andrew	References, internal feedback	--
0.3	Draft	12/05/2021 / Black Andrew	Internal feedback	--
1.0	Approved	17/12/2021 / Black Andrew	First release	HP
1.1	Approved	07/02/2022 / Black Andrew	Feedback from users	HP
1.2	Approved	06/05/2022 / Black Andrew	Feedback from users	HP
1.3	Approved	27/06/2022 / Black Andrew	Minor typos, feedback from users	HP

Table of Contents

WindCube Data Filtering Guidelines

1 Executive Summary

Filter Type	10-Minute Availability	CNR
Highest Speed Precision	90%	> -20 dB
WindCube v2 and v2.1 IEC 61400-12-1 Classification Filter Thresholds	80%	None
Highest Campaign Availability	50%	None
	Or use Uncertainty Tables	

Table 1: Vaisala Manufacturer Recommended Filtering Guidelines for WindCube

These filters are Vaisala Manufacturer Recommendations for WindCube v2 and v2.1. For energy yield assessment, filters should be applied monthly. To meet Monthly Campaign Data Availability requirements, the filter thresholds can be progressively loosened to successively higher uncertainty (Availability, CNR) bins.

The values of 10-minute Data Availability and 10-minute average CNR can be found in WindCube .STA files downloaded from WindWeb, or accessed directly from the device with column names:

- `XXm CNR (dB)`
- `XXm Data Availability (%)`

where XX corresponds to the measurement height (e.g. 40, 80, 100, 150, 300)

In principle, to minimize the uncertainties in estimates of long-term wind speed statistics, it is preferable to include as many 10-minute samples as possible and to estimate the associated uncertainties rather than filtering data and leaving gaps with no wind information.

2 Purpose

WindCube profiling lidars (v1, v2, v2.1) are used for wind energy and meteorological applications. This document describes the Vaisala's recommended data filtering guidelines for the WindCube profiler product line.

Remote sensing devices such as lidar, radar, and sodar rely on atmospheric backscatter targets to reflect energy back to the sensor to gather information about the wind. The atmosphere is a highly variable medium, and the concentration of these backscatter targets can vary at a variety of timescales from diurnally to seasonally. In the case of lidar, the primary backscatter targets are aerosols, large molecules, suspended in the air, advecting with the wind.

Variability in backscatter target concentration causes variability in the quality and quantity of measurement samples of the wind speed by the WindCube within a typical 10-minute measurement period. The uncertainty of individual line-of-sight (LOS) measurements varies with aerosol concentration, and the exclusion of LOS measurements below a fixed uncertainty threshold causes variability in the number of measurements averaged together within a typical 10-minute measurement period.

The 10-minute uncertainty of WindCube measurements can be traced to these two parameters: (1) the average Carrier-to-Noise Ratio (CNR), or signal quality, of the LOS measurements, and (2) the number of samples above the minimum CNR threshold. The WindCube measurement uncertainty can be assessed with three different methods: (1) calculating a minimum uncertainty using laboratory tests of LOS and CNR, (2) calculating a maximum uncertainty using 400 factory validation tests with "Golden" WindCube and (3) unit specific uncertainty using on-site mast verification (calibration) process following IEC 61400-12-1 ed2 Annex L.3. In this document, we estimate the minimum and maximum uncertainties for a WindCube lidar as a function of 10-minute average CNR and number of samples in a 10-minute period. The minimum uncertainties are estimated using laboratory tests of WindCube LOS uncertainty as a function of CNR. For all possible CNRs and sample numbers, these uncertainties are propagated through the WindCube's Wind Field Reconstruction (WFR) algorithm to estimate the 10-minute measurement uncertainty. The maximum uncertainties are estimated using factory validation test data from more than 400 WindCubes. Each WindCube collected data simultaneously with a "Golden" WindCube, separated on the factory roof by no more than 4 meters. Treating these lidars as identical, collocated instruments, we estimate the maximum uncertainty of the WindCube data as a function of CNR and number of samples via the distributions of residual standard deviation.

Our goal in estimating minimum and maximum uncertainties for WindCube-measured wind speeds is to enable WindCube users to use as much data as possible for their applications and to understand the associated uncertainties of the data. **In principle, to minimize the uncertainties in estimates of long-term wind speed statistics, it is preferable to include as many 10-minute samples as possible and to estimate the associated uncertainties rather than filtering data and leaving gaps, with no wind information.** This document attempts to define the limits of this approach to wind measurement for the WindCube.

3 Definitions and References

3.1 Carrier-to-Noise Ratio (CNR)

The CNR is defined as the ratio between the detected signal power and the wideband noise power in the Doppler spectrum. WindCube v2 and v2.1 use a binary filter threshold of -23 dB, excluding any 1 Hz horizontal wind speed measurement containing a radial Line-of-Sight (LOS) measurement with CNR < -23 dB.

3.2 Data Availability

3.2.1 10-minute availability

The scalar 10-minute availability is defined, at each altitude:

$$100 * \frac{Count(1\ Hz\ horizontal\ wind\ speeds\ with\ status\ =\ 1)}{600}$$

Equation 1: Scalar WFR 10-minute Data Availability

This value is reported for each 10-minute wind speed, for each altitude.

The vector 10-minute availability is defined, at each altitude:

$$100 * \frac{Count(LOS\ wind\ speeds\ with\ status\ =\ 1)}{600}$$

Equation 2: Vector WFR 10-minute Data Availability

This value is reported for each 10-minute wind speed, for each altitude.

3.2.2 Campaign availability

The Campaign Availability is defined, at each altitude:

$$100 * \frac{Count(10\ minute\ wind\ speeds\ meeting\ filter\ criteria)}{Count(10\ minute\ intervals\ in\ measurement\ campaign)}$$

Equation 3: Overall Campaign Availability

Typical Campaign Availability requirements for wind resource assessment are 80% or 90%. Check with your campaign stakeholders to determine which availability requirement should be followed.

See MEASNET Edition 2 (2016) Chapter 7.2 Wind Speed, Page 12.

3.3 Line-of-Sight (LOS)

Line-of-Sight, LOS, and radial wind speed (RWS) all refer to wind speeds measured along the WindCube's oblique 28° beams. There are 5 beams, North, South, East, West, and Vertical which are combined via Wind Field Reconstruction to compute the Cartesian wind speed components: horizontal wind speed, vertical wind speed (simply the vertical LOS), wind direction, and turbulence intensity.

3.4 Wind Field Reconstruction (WFR)

Wind Field Reconstruction is a class of algorithms used to convert LOS wind speeds to Cartesian wind speeds.

3.4.1 Scalar WFR

In Scalar WFR, the latest four (NSEW) LOS wind speeds are combined to generate a horizontal wind speed measurement at 1Hz sampling rate. At the end of the 10-minute period, these 600 horizontal wind speeds are averaged together.

Scalar WFR is used in WindCube v2

There must be 4 consecutive LOS measurements to generate a horizontal wind speed estimate:

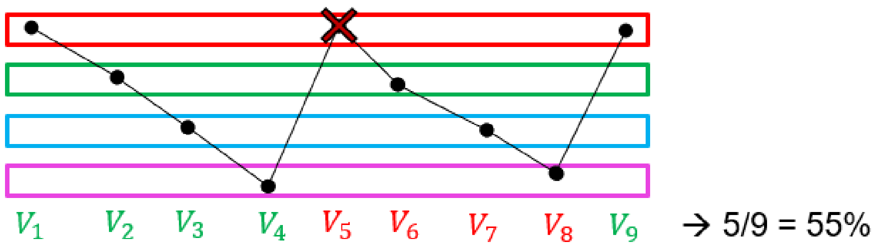

Figure 1: Illustration of Scalar WFR and Data Availability. X corresponds to an LOS with CNR < -23 dB

$$u_i = \frac{V_{r0} - V_{r180}}{2\ sin(\theta)}$$

$$v_i = \frac{V_{r270} - V_{r90}}{2\ sin(\theta)}$$

$$\overline{V_{scalar}} = \frac{1}{600}\sum_{i=1}^{600}\sqrt{u_i^2 + v_i^2}$$

Equation 4: Scalar Wind Field Reconstruction Equations

3.4.2 Vector WFR

In Vector WFR, the 150 LOS wind speeds are individually averaged at the end of the 10-minute period, and then combined to generate the horizontal wind speed. There is only one Vector WFR wind speed per 10-minute period.

Vector WFR 10-minute availability is computed based on the total LOS availability:

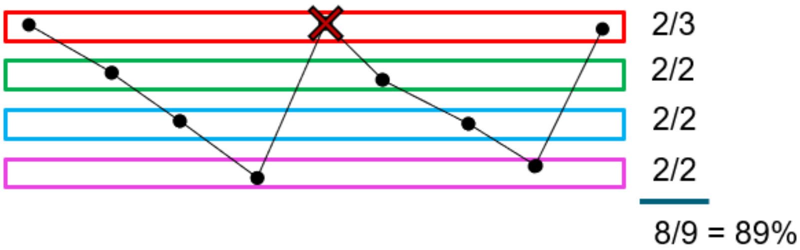

Figure 2: Illustration of Vector WFR and Data Availability

$$\bar{u} = \frac{\overline{V_{r0}} - \overline{V_{r180}}}{2\,sin(\theta)}$$

$$\bar{v} = \frac{\overline{V_{r270}} - \overline{V_{r90}}}{2\,sin(\theta)}$$

$$\overline{V_{vector}} = \sqrt{\bar{u}^2 + \bar{v}^2}$$

Equation 5: Vector Wind Field Reconstruction Equations. u and v same as above

3.4.3 Hybrid WFR

In Hybrid WFR, the Scalar and Vector WFR wind speeds are combined:

$$V_{Hybrid}^{lidar} = \frac{2}{3}V_{scalar}^{lidar} + \frac{1}{3}V_{vector}^{lidar}$$

Equation 6: Hybrid wind field reconstruction weightings

Hybrid WFR is used in WindCube v2.1

The 10-minute availability reported for WindCube v2.1 is from Scalar WFR

4 Estimates of WindCube 10-Minute Uncertainty

4.1 Minimum Uncertainties

The minimum uncertainty of 10-minute WindCube measurements is derived from laboratory tests of the WindCube's LOS uncertainty, propagated through the wind field reconstruction algorithm.

4.1.1 Laboratory LOS Uncertainty Estimates

The laboratory tests are carried out using a test apparatus named Simulation of the Atmosphere using Fiber Optics (SAFO). SAFO creates a controlled propagation medium, with motionless scattering particles. The scattering medium mimics the atmospheric-distributed target behavior, as well as the speckle, without velocity gradient or true atmospheric turbulence. Velocity is estimated on 10 successive range gates, each having their own CNR value. Repeated measurements are performed to estimate the mean speed and the speed deviation. Comparisons between these three techniques show good agreement and show consensus on the behavior of δV_r as a function of CNR.

4.1.2 Estimate of Minimum Uncertainty

For each SAFO measurement of LOS uncertainty as a function of CNR, we propagate δV_r through the WFR for various sample numbers and CNRs (see Annex)

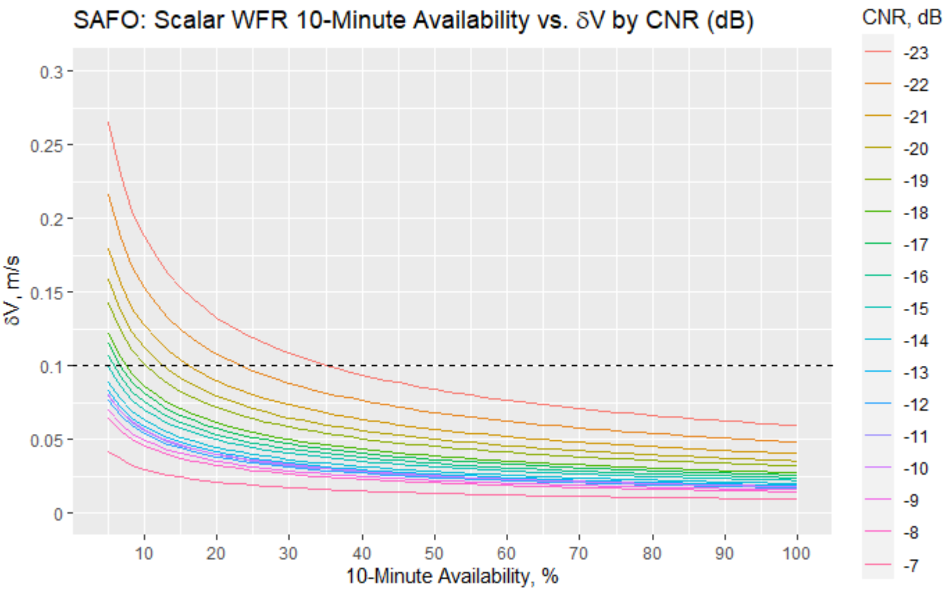

Figure 3: Minimum WindCube uncertainty estimated using SAFO and Vector WFR, various CNR and Availability. Dashed line corresponds to 0.1 m/s, the specified precision of the WindCube

4.2 Maximum Uncertainties

To estimate maximum uncertainties as a function of CNR and Data Availability, we use the manufacturing validation data from 448 WindCube v2s (referred to as unit under test or UUT), and fit a function derived from the Cramer-Rao lower bound estimate of the lidar uncertainty, and the WFR algorithm

4.2.1 Manufacturing Test Uncertainty Estimates

Some devices were tested more than once, after regular maintenance cycles. There were three different Golden WindCubes used in these validations, each of which was validated against an IEC-compliant met mast before and after use in factory tests, at 2-year intervals.

Any devices that fail the Vaisala validation test criteria are excluded from this analysis, as these units would not leave the factory without intervention.

Both the Golden Lidar and the test unit were binned by CNR and Availability into the groups listed in the Annex. For the purposes of this analysis, the Golden and UUT are considered two identical instruments. In each bin, the experimental standard deviation of the mean is computed. This data was fit to an approximation of the Cramer-Rao lower bound (CNR) and the wind field reconstruction algorithm (Availability).

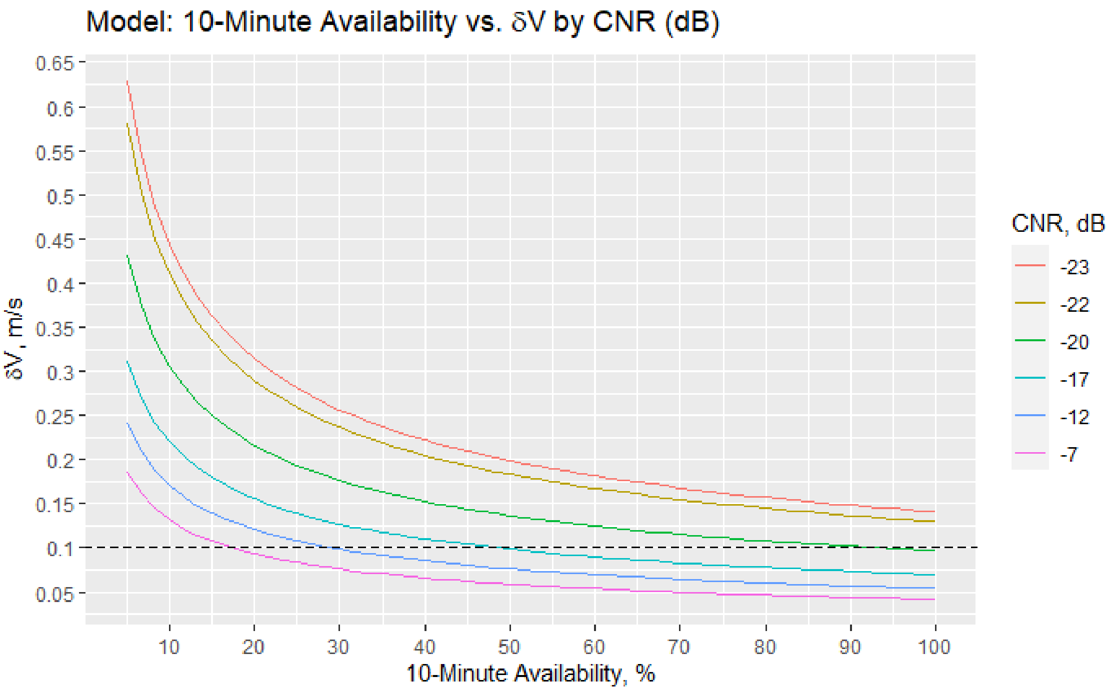

Figure 4: Estimate of maximum wind speed uncertainty as a function of CNR and 10-minute Data Availability

This model has the same functional form as the SAFO+WFR data shown in the previous section, with slightly increased magnitude. We expect this increase in magnitude due to two effects: the offset in space of the two lidars of a few meters; and offsets in time due to LOS beam asynchronization.

4.3 Uncertainty Ranges

The estimates of minimum and maximum uncertainties are well behaved. Both show characteristic curves following $1/\sqrt{N}$ law with respect to Availability and 1/CNR with respect to CNR.

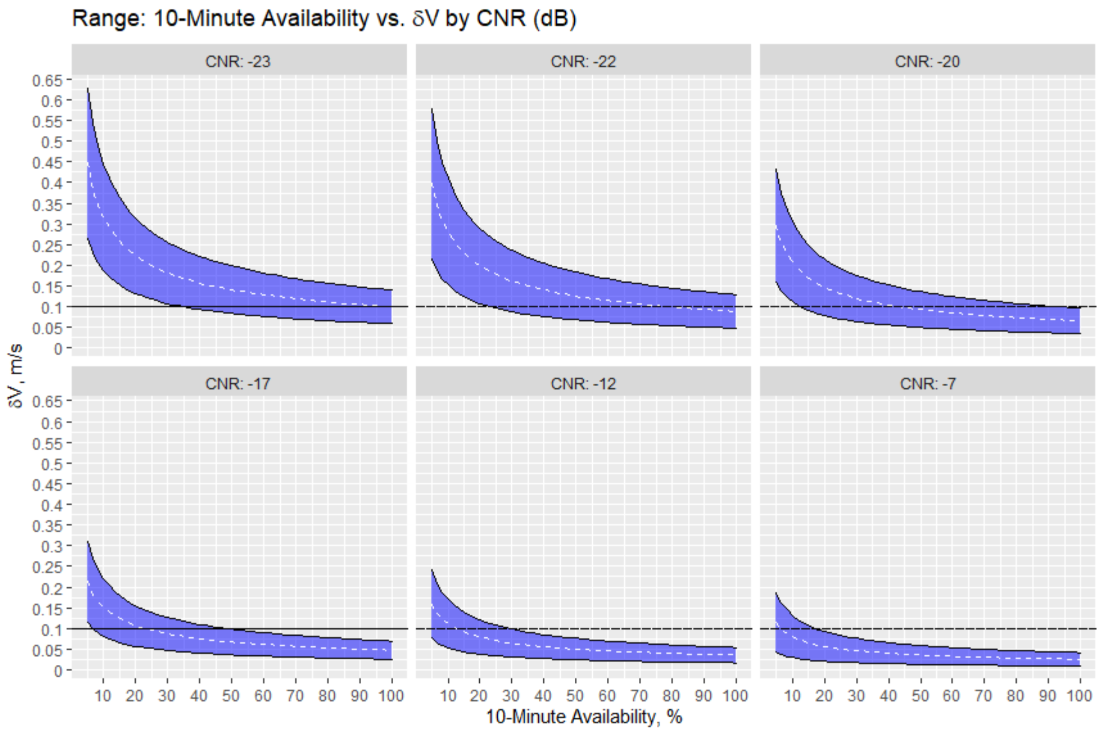

Figure 5: Estimates of minimum and maximum wind speed uncertainty by availability and CNR

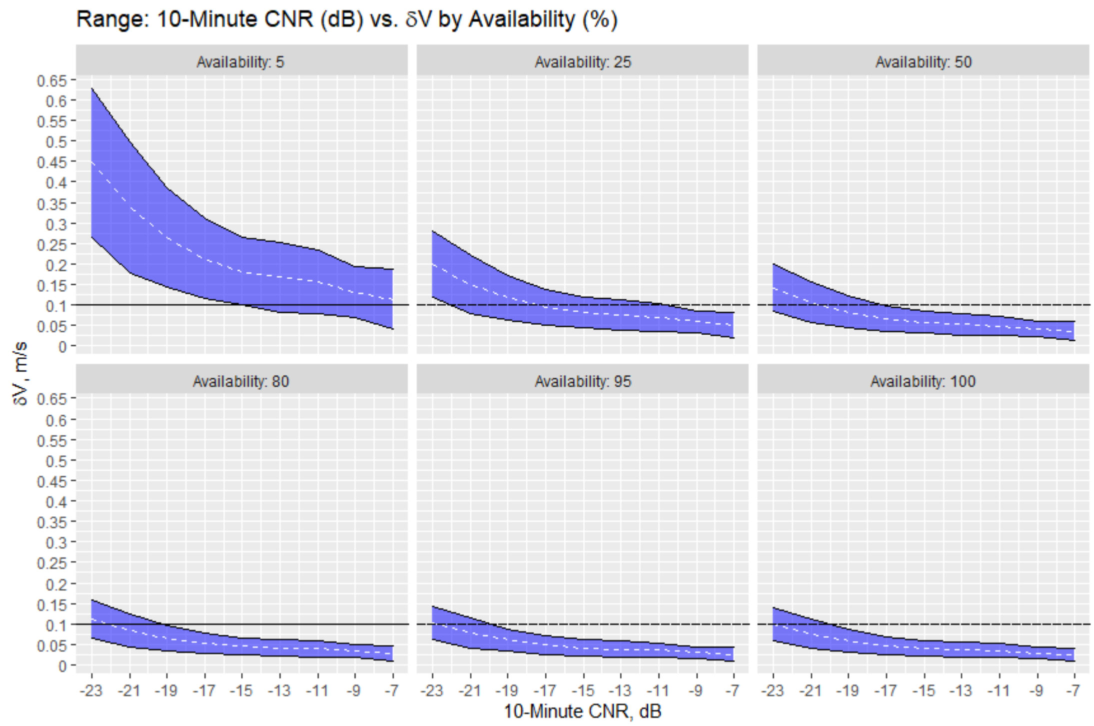

Figure 6: Estimates of minimum and maximum wind speed uncertainty by CNR and availability

4.4 Uncertainty Tables

The minimum, average, and maximum estimates of WindCube precision in each Availability decile and CNR decibel are shared here. These may be used to develop estimates of overall WindCube uncertainties. Note that these do not incorporate sensitivities to Environmental Parameters, as documented in IEC 61400-12-1 style sensitivity test.

4.4.1 Minimum estimate of WindCube precision

CNR (dB) \ Avail (%)	0	10	20	30	40	50	60	70	80	90	100
-23	0.446	0.156	0.119	0.100	0.088	0.080	0.074	0.068	0.064	0.061	0.059
-22	0.363	0.127	0.097	0.082	0.072	0.065	0.060	0.056	0.052	0.050	0.048
-21	0.302	0.106	0.081	0.068	0.060	0.054	0.050	0.046	0.044	0.041	0.040
-20	0.266	0.093	0.071	0.060	0.053	0.048	0.044	0.041	0.038	0.036	0.035
-19	0.240	0.084	0.064	0.054	0.048	0.043	0.040	0.037	0.035	0.033	0.032
-18	0.205	0.072	0.055	0.046	0.041	0.037	0.034	0.032	0.030	0.028	0.027
-17	0.193	0.068	0.052	0.044	0.038	0.035	0.032	0.030	0.028	0.026	0.026
-16	0.179	0.063	0.048	0.040	0.036	0.032	0.030	0.027	0.026	0.024	0.024
-15	0.167	0.058	0.045	0.038	0.033	0.030	0.028	0.026	0.024	0.023	0.022
-14	0.150	0.052	0.040	0.034	0.030	0.027	0.025	0.023	0.022	0.020	0.020
-13	0.140	0.049	0.037	0.032	0.028	0.025	0.023	0.021	0.020	0.019	0.019
-12	0.128	0.045	0.034	0.029	0.025	0.023	0.021	0.020	0.018	0.017	0.017
-11	0.133	0.047	0.036	0.030	0.026	0.024	0.022	0.020	0.019	0.018	0.018
-10	0.135	0.047	0.036	0.030	0.027	0.024	0.022	0.021	0.019	0.018	0.018
-9	0.118	0.041	0.032	0.027	0.023	0.021	0.019	0.018	0.017	0.016	0.016
-8	0.108	0.038	0.029	0.024	0.022	0.019	0.018	0.017	0.016	0.015	0.014
≥ -7	0.071	0.025	0.019	0.016	0.014	0.013	0.012	0.011	0.010	0.010	0.009

Table 1: Minimum estimates of WindCube precision as a function of CNR and Availability. All figures m/s. Bin labels represent lower bin edges (i.e. Avail column 90 covers Data Availability values 90% - 100%)

4.4.2 Average estimate of WindCube precision

CNR (dB) \ Avail (%)	0	10	20	30	40	50	60	70	80	90	100
-23	0.553	0.263	0.201	0.169	0.149	0.135	0.124	0.115	0.108	0.103	0.100
-22	0.486	0.234	0.179	0.151	0.133	0.120	0.110	0.103	0.097	0.091	0.089
-21	0.413	0.199	0.152	0.128	0.113	0.102	0.094	0.087	0.082	0.078	0.076
-20	0.360	0.174	0.133	0.112	0.099	0.089	0.082	0.076	0.072	0.068	0.066
-19	0.323	0.156	0.119	0.100	0.088	0.080	0.073	0.068	0.064	0.061	0.059
-18	0.284	0.137	0.105	0.088	0.078	0.070	0.065	0.060	0.057	0.054	0.052
-17	0.260	0.125	0.096	0.081	0.071	0.064	0.059	0.055	0.052	0.049	0.048
-16	0.239	0.115	0.088	0.074	0.065	0.059	0.054	0.051	0.048	0.045	0.044
-15	0.222	0.107	0.082	0.069	0.061	0.055	0.050	0.047	0.044	0.042	0.041
-14	0.214	0.104	0.080	0.067	0.059	0.053	0.049	0.046	0.043	0.041	0.040
-13	0.202	0.099	0.075	0.063	0.056	0.051	0.046	0.043	0.041	0.038	0.037
-12	0.191	0.093	0.071	0.060	0.053	0.048	0.044	0.041	0.039	0.036	0.035
-11	0.189	0.092	0.070	0.059	0.052	0.047	0.043	0.040	0.038	0.036	0.035
-10	0.186	0.090	0.069	0.058	0.051	0.046	0.043	0.040	0.037	0.035	0.034
-9	0.161	0.078	0.059	0.050	0.044	0.040	0.037	0.034	0.032	0.030	0.029
-8	0.154	0.075	0.057	0.048	0.042	0.038	0.035	0.033	0.031	0.029	0.028
≥ -7	0.133	0.067	0.051	0.043	0.038	0.034	0.032	0.030	0.028	0.026	0.026

Table 2: Average estimates of WindCube precision as a function of CNR and Availability. All figures m/s. Bin labels represent lower bin edges (i.e. Avail column 90 covers Data Availability values 90% - 100%)

4.4.3 Maximum estimate of WindCube precision

CNR (dB) \ Avail (%)	0	10	20	30	40	50	60	70	80	90	100
-23	0.661	0.370	0.283	0.238	0.210	0.190	0.175	0.162	0.153	0.144	0.141
-22	0.609	0.341	0.261	0.220	0.194	0.175	0.161	0.150	0.141	0.133	0.130
-21	0.523	0.293	0.224	0.189	0.166	0.150	0.138	0.129	0.121	0.114	0.111
-20	0.454	0.254	0.195	0.164	0.144	0.130	0.120	0.112	0.105	0.099	0.097
-19	0.405	0.227	0.174	0.146	0.129	0.116	0.107	0.100	0.094	0.088	0.086
-18	0.362	0.203	0.155	0.131	0.115	0.104	0.096	0.089	0.084	0.079	0.077
-17	0.327	0.183	0.140	0.118	0.104	0.094	0.086	0.080	0.076	0.071	0.070
-16	0.300	0.168	0.128	0.108	0.095	0.086	0.079	0.074	0.069	0.065	0.064
-15	0.278	0.156	0.119	0.100	0.088	0.080	0.073	0.068	0.064	0.061	0.059
-14	0.278	0.156	0.119	0.100	0.088	0.080	0.073	0.068	0.064	0.061	0.059
-13	0.264	0.148	0.113	0.095	0.084	0.076	0.070	0.065	0.061	0.058	0.056
-12	0.253	0.142	0.109	0.091	0.081	0.073	0.067	0.062	0.059	0.055	0.054
-11	0.245	0.137	0.105	0.088	0.078	0.070	0.065	0.060	0.057	0.053	0.052
-10	0.238	0.133	0.102	0.086	0.076	0.068	0.063	0.059	0.055	0.052	0.051
-9	0.204	0.114	0.087	0.073	0.065	0.058	0.054	0.050	0.047	0.044	0.043
-8	0.199	0.112	0.085	0.072	0.063	0.057	0.053	0.049	0.046	0.044	0.042
≥ -7	0.196	0.110	0.084	0.071	0.062	0.056	0.052	0.048	0.045	0.043	0.042

Table 3: Maximum estimates of WindCube precision as a function of CNR and Data Availability. All figures m/s. Bin labels represent lower bin edges (i.e. Avail column 90 covers Data Availability values 90% - 100%)

4.4.4 Additional Uncertainty following IEC 61400-12-1

In cases where campaigns comply with the IEC 61400-12-1 standard, the filtering thresholds are most simply the same as those used when the device classification was carried out.

In cases where those IEC Classification filters (Section 4.4, Table 3, WindCube v2 and v2.1 IEC 61400-12-1 Classification Filter Thresholds) lead to unacceptably low campaign availability, it is permissible to loosen the filter thresholds as described in the standard:

Annex L.2.3 Data preparation, Section (e)

> *There is a possibility that the data availability requirements may be relaxed. The influence that relaxing these requirements could have on the measurement accuracy should be checked by means of a sensitivity analysis.*

Annex L.2.4 Principle and requirements of a sensitivity test

> *The initial sensitivity to be tested shall relate to the relaxation of the criterion described in item e) in L.2.3. The sensitivity of the deviation between the reference sensors and the remote sensing device on the availability of the remote sensing device within the averaging interval or the corresponding quality factor of the measurements (as defined by the manufacturer) shall be examined. Remote sensing data shall be deemed acceptable if characterised by a level of availability or by a quality factor that has been demonstrated to have no significant influence on the deviation between the reference sensors and the remote sensing device measurements.*

In the Classification tests for WindCube v2 and v2.1 Availability was not identified as a significant Environmental Parameters. The Classification tests included data with Data Availability between 80% and 100% and various CNRs.

The data in the Annex, (8.1 Manufacturing Tests Against Golden Lidar 2012 – 2017) and the theories developed in this section show that the precision of the WindCube varies as a function of CNR and Availability. There is not a physical mechanism that would cause data with lower Availability to respond differently to Environmental Parameters such as wind shear, turbulence intensity, wind veer, or other atmospheric phenomena with respect to bias. There is, however, a clear, demonstrable increase in the uncertainty of the device due to (1) the increase in LOS uncertainty as a function of CNR, and (2) the decrease in sample number as more LOS measurements' CNRs fall below the -23 dB cutoff threshold, as shown in the error propagation through WFR.

Error propagation of LOS uncertainty through WFR is documented in the annex and can be used to add an additional uncertainty term in those cases with data filtering that deviates from the Classification test.

The additional uncertainties in the table below are computed:

$$\Delta\sigma_{RSD,(Avail,CNR)} = \sqrt{\sigma^2_{RSD,(Avail,CNR)} - \sigma^2_{RSD,(80\%,CNR)}}$$

for each CNR and Availability bin. The terms below can be added in quadrature to existing IEC 61400-12-1 uncertainty expressions.

CNR (dB) \ Avail (%)	0	10	20	30	40	50	60	70
-23	0.642	0.335	0.235	0.179	0.139	0.106	0.076	0.041
-22	0.592	0.309	0.217	0.165	0.128	0.098	0.070	0.038
-21	0.508	0.265	0.186	0.142	0.110	0.084	0.060	0.032
-20	0.441	0.230	0.162	0.123	0.096	0.073	0.052	0.028
-19	0.393	0.206	0.144	0.110	0.085	0.065	0.047	0.025
-18	0.351	0.183	0.129	0.098	0.076	0.058	0.042	0.022
-17	0.318	0.166	0.117	0.089	0.069	0.053	0.038	0.020
-16	0.291	0.152	0.107	0.081	0.063	0.048	0.034	0.019
-15	0.270	0.141	0.099	0.075	0.059	0.045	0.032	0.017
-14	0.270	0.141	0.099	0.075	0.059	0.045	0.032	0.017
-13	0.257	0.134	0.094	0.072	0.056	0.043	0.030	0.016
-12	0.246	0.129	0.090	0.069	0.053	0.041	0.029	0.016
-11	0.238	0.124	0.087	0.066	0.052	0.039	0.028	0.015
-10	0.231	0.121	0.085	0.065	0.050	0.038	0.027	0.015
-9	0.198	0.103	0.073	0.055	0.043	0.033	0.023	0.013
-8	0.194	0.101	0.071	0.054	0.042	0.032	0.023	0.012
>-7	0.190	0.099	0.070	0.053	0.041	0.032	0.022	0.012

Table 4: Additional uncertainty, $\Delta\sigma_{RSD,(Avail,CNR)}$, to use in IEC 61400-12-1 as function of CNR and Data Availability. All figures m/s. Bin labels represent lower bin edges (i.e. Avail column 90 covers Data Availability values 90% - 100%)

Note that in each Availability decile, the uncertainty values are the average of 1% Availability bins, and thus are slightly lower than the differences in quadrature taken from Table 3.

4.4.5 Worked Example

In a case where Monthly Campaign Data Availability (MCDA) is 50%, but the project requires 80% MCDA, the filter thresholds can be loosened to add the minimum additional uncertainty, while meeting the MCDA requirement, and complying with ICE 61400-12-1 as closely as possible.

Maximum monthly samples:	24 * 6 * 30 = 4320
80% MCDA samples requirement:	0.8 * 4320 = 3456
Samples, Availability > 80%:	0.5 * 4320 = 2160
	(1296 more needed for MCDA requirement)

Examine the dataset, and select bins containing data with lowest uncertainty

Samples. 70% < Availability < 80%, CNR = -21 dB:	801
Samples, 70% < Availability < 80%, CNR = -22 dB:	522

The samples in the (70%, -21 dB) and (70%, -22 dB) bins are added to the dataset.

Total samples:	3483
New MCDA:	80.6% (meets requirement)
Percentage of dataset in (70%, -21 dB):	23% (801 / 3483)
Percentage of dataset in (70%, -22 dB):	15% (522 / 3483)

Additional uncertainty terms are added in quadrature to the existing uncertainties, in m/s:

$$\sigma_{RSD,(Avail,CNR)} = \sum_{i=(Avail,CNR)}^{N} \frac{M_i}{S} {\sigma_{RSD,i}}^2$$

Where M is the number of samples in the (*Avail, CNR*) bin, and S is the total number of samples in the expanded dataset.

$$\sigma_{RSD} = \sqrt{\sigma_{RSD,>80\%} + \sigma_{RSD,(Avail,CNR)}}$$

$$\sigma_{RSD} = \sqrt{\sigma_{RSD,>80\%} + \frac{M_1}{S} {\sigma_{RSD,1}}^2 + \frac{M_2}{S} * {\sigma_{RSD,2}}^2}$$

Where $\sigma_{RSD,>80\%}$ includes all uncertainties associated with IEC Classification and Verification.

With the Example data, and values taken from Table 4:

$$\sigma_{RSD} = \sqrt{\sigma_{RSD,>80\%} + \cdots + 0.23 * 0.032^2 + 0.15 * 0.038^2} = \sqrt{\sigma_{RSD,>80\%} + 0.0004}$$

To add this data as a percentage, the same is carried out but in each wind speed bin, normalized by the wind speed:

$$\sigma_{RSD,3ms^{-1}} = \sqrt{\sigma_{RSD,>80\%,3ms^{-1}} + \cdots + 0.23 * \left(\frac{0.032}{3.5}\right)^2 + 0.15 * \left(\frac{0.038}{3.5}\right)^2} = \sqrt{\sigma_{RSD,>80\%} + 0.000037}$$

Examining only the (Avail, CNR) bin terms themselves, in percentage:

$$100 * \sqrt{0.23 * \left(\frac{0.032}{3.5}\right)^2 + 0.15 * \left(\frac{0.038}{3.5}\right)^2} = 0.61\%$$

This shows a +0.61% increase in uncertainty in the 3 m/s – 4 m/s bin when 23% and 15% of the data come from the (70%, -21 dB) and (70%, -22 dB) bins.

5 Data Filtering Thresholds

The subsections below describe Data Filtering Thresholds for WindCube horizontal wind speeds. There are two primary wind energy applications for WindCube data: energy yield assessment and power performance testing. These applications have different demands on data filtering.

Energy Yield Assessment

In energy yield assessment, high monthly horizontal wind speed data availability is critical for bankable measurement campaigns. **Typical requirements for monthly campaign availability are 80% or 90%**. Verify and document the target campaign availability requirements while planning your measurement campaign.

Power Performance Testing

In contractual power performance testing using lidar, the accuracy of the device is more important than data availability. To explicitly follow the terms of the IEC 61400-12-1 standard, one must attempt to use the **same filters as during the remote sensors' IEC Classification**. Exceptions to this must be carefully documented and should reference the uncertainty tables contained in this document. Consult with and agree upon data filtering thresholds with all PPT measurement campaign stakeholders.

5.1 10-Minute Availability

There are three levels of horizontal wind speed filtering possible based on the 10-Minute Availability reported with each wind speed.

Highest Speed Precision: 90%

Using a threshold of 90% 10-Minute Availability yields data with precision limit of 0.1 m/s than lower thresholds, while giving Campaign Availability > 80% in most circumstances. This threshold is based on the data in Table 3.

WindCube v2 and v2.1 IEC 61400-12-1 Classification Filter Thresholds: 80%

Vaisala's recommended practice for data filtering in power performance testing is to use a threshold of 80% for 10-Minute Availability, expanding the thresholds, if necessary, using the uncertainties in Table 4.

WindCube v2 and v2.1 IEC Classification Reports use 80% 10-Minute Availability threshold.

Highest Campaign Availability: 50% (or Use Uncertainty Tables)

In circumstances where Campaign Availability falls below the desired percentage on a monthly or campaign basis, the 10-Minute Availability threshold can be lowered incrementally following the Uncertainty Tables to increase Campaign Availability up to 80% or 90% as required by the project stakeholders. Document the percentage of data in each bin of the Uncertainty Table, and which table used. This threshold is based on the data in Table 1, with a small safety factor added.

5.2 CNR

There are three levels of horizontal wind speed filtering possible based on the CNR reported with each wind speed.

Highest Speed Precision: -20 dB

WindCube v2 and v2.1 IEC 61400-12-1 Classification Filter Thresholds: None

Highest Campaign Availability: None (or Use Uncertainty Table)

5.3 Summary

<table>
<tr><th>Filter Type</th><th>10-Minute Availability</th><th>CNR</th></tr>
<tr><td>Highest Speed Precision</td><td>90%</td><td>> -20 dB</td></tr>
<tr><td>WindCube v2 and v2.1
IEC 61400-12-1 Classification Filter Thresholds</td><td>80%</td><td>None</td></tr>
<tr><td rowspan="2">Highest Campaign Availability</td><td>50%</td><td>None</td></tr>
<tr><td colspan="2">*Or use Uncertainty Tables*</td></tr>
</table>

Table 5: Vaisala Manufacturer's Recommended Filtering Guidelines for WindCube v2 and v2.1

These filters are Vaisala Manufacturer Recommendations for WindCube v2 and v2.1. For energy yield assessment, filters should be applied monthly. To meet Monthly Campaign Data Availability requirements, the filter thresholds can be progressively loosened to successively higher uncertainty (Avail, CNR) bins

6 Conclusions

The WindCube's precision and accuracy are well understood and closely follow heterodyne lidar measurement theory and statistical sampling theory. WindCube users may refer to this guideline, as well as the references when using WindCube data for wind energy applications including bankable energy yield assessment, contractual power performance testing, and any other application demanding traceable precision or accuracy reporting.

The relationships between CNR, 10-minute availability and wind speed precision demonstrated in this Guideline should give WindCube users confidence that they may adjust filter criteria between the Highest Campaign Availability and Highest Speed Precision recommendations to meet their measurement campaign requirements. The precision sensitivity between these two thresholds, along the CNR and 10-minute availability axes, is well-behaved, and follows from the physical principles of lidar measurement.

7 Annex A: Algorithms

7.1 Uncertainty Propagation through Scalar WFR

We propagate the LOS uncertainties through the WFR algorithm:

$$\delta u^2 = \left[\frac{1}{2\sin\theta}\right]^2 \left(\delta V_{r0}{}^2 + \delta V_{r180}{}^2\right)$$

$$\delta u = \delta v = \frac{1}{2\sin\theta}\sqrt{2}\,\delta V_r$$

$$Vh^2 = u^2 + v^2 \rightarrow \delta Vh^2 = Vh\frac{\sqrt{2}\,\delta V_r}{\sin\theta}$$

$$\delta Vh^2 = Vh\frac{\sqrt{2}\,\delta V_r}{\sin\theta}$$

$$\delta Vh = \frac{\sqrt{2}\,\delta V_r}{2\sin\theta}$$

Summing this uncertainty to 10-minute data, we have the equation:

$$\boldsymbol{\delta Vh_{10min}} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sqrt{\left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2 + cov\left[\delta V_{r,i,j}\right]}$$

Equation 7: Uncertainty of 10-minute scalar wind speed as function of LOS uncertainty and availability

$cov\left[\delta V_{r,i,j}\right]$ is the covariance between individual 1 Hz measurements. Note that this covariance term need not reflect covariance in the wind, but rather the covariance in LOS uncertainties common between successive 1 Hz. Each 1 Hz scalar wind speed shares 3 LOS with the sample n±1, 2 LOS with n±2, and 1 LOS with n±3. Taking the convolution of two length-4 moving-average boxcar windows, and excluding the central point, the default value for the covariance is:

$$cov\left[\delta V_{r,i,j}\right] = 2 * \left(0.75 * \left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2 + 0.5 * \left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2 + 0.25 * \left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2\right) = 3 * \left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2$$

Equation 8: Covariance term for Scalar WFR error propagation

N is the number of samples; δVh_{10min} is the standard error for horizontal wind speed measurements; θ is the WindCube's beam tilt angle (28°). We assume δV_r is uniform for all four beams. Uncertainties in the beam angle are omitted as they are an order of magnitude smaller than the LOS uncertainties.

$$\delta Vh_{10min} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sqrt{4 * \left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2} = \frac{2}{\sqrt{N}}\sum_{i=1}^{N}\sqrt{\left(\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta}\right)^2} = \frac{1}{\sqrt{N/4}}\sum_{i=1}^{N}\frac{\sqrt{2}\,\delta V_{r,i}}{2\sin\theta} = \frac{\sqrt{2}\,\delta V_r}{2\sqrt{N_r}\sin\theta}$$

Equation 9: Simplified Scalar WFR error propagation

7.2 Uncertainty Propagation through Vector WFR

For vector wind field reconstruction, the LOS speeds are summed before reconstruction, so we combine the uncertainties earlier:

$$\overline{V_{r0}} = \frac{1}{N_{r0}} \sum_{i=1}^{N_{r0}} V_{r0,i} \rightarrow \delta V_{r0} = \frac{1}{N_{r0}} \sqrt{\sum_{i=1}^{N_{r0}} \delta V_{r0,i}{}^{2}}$$

$$\delta V_{r0} = \frac{1}{N_{r0}} \sqrt{N_{r0} \delta V_{r0}{}^{2}} = \frac{\delta V_{r0}}{\sqrt{N_{r0}}}$$

Now we follow the same formula for a 1 Hz scalar WFR using this average LOS uncertainty to get the vector WFR 10-minute uncertainty:

$$\delta u = \delta v = \frac{1}{2 \sin \theta} \frac{\sqrt{2}}{\sqrt{N_r}} \delta V_r$$

$$Vh^2 = u^2 + v^2 \rightarrow \ \delta Vh^2 = \ Vh \frac{\sqrt{2}\, \delta V_r}{\sqrt{N_r} \sin \theta}$$

$$\delta Vh^2 = \ Vh \frac{\sqrt{2}\, \delta V_r}{\sqrt{N_r} \sin \theta}$$

$$\boldsymbol{\delta V h_{10min}} = \frac{\boldsymbol{\sqrt{2}\, \delta V_r}}{\boldsymbol{2 \sqrt{N_r} \sin \theta}}$$

Equation 10: Uncertainty of 10-minute vector WFR wind speed as a function of LOS uncertainty and availability

Note that the scalar and vector WFR 10-minute uncertainties are identical for the same Availability.

7.3 Experimental Standard Deviation of the Mean

For each campaign, the residual standard deviation is computed:

$$\sigma(V_{Golden} - V_{WC}) = \sqrt{\frac{\sum_{j=1}^{n} \left[\left(\overline{V_{Golden} - V_{WC}} \right) - \left(V_{Golden,j} - V_{WC,j} \right) \right]^2}{n - 1}}$$

or more compactly:

$$\sigma(\epsilon_k) = \sqrt{\frac{\sum_{j=1}^{k} \left(\bar{\epsilon} - \epsilon_j \right)^2}{n - 1}}$$

Where $\epsilon_k = V_{Golden} - V_{WC}$, are the *residuals* of the validation data. Taking the standard deviation of the residuals removes any offsets between the two devices that may be present from variation in components and focuses on the device precision as a function of CNR and Availability.

The two devices contribute to the overall standard deviation:

$$\sigma = \sqrt{\sigma_{Golden}^2 + \sigma_{WC}^2}$$

If we assume that the uncertainty of the Golden WindCube is roughly the same as a typical WindCube:

$$v_{Golden} \approx v_{WC} \Rightarrow \sigma = \sqrt{2\sigma_{WC}^2} \therefore \sigma_{WC} = \frac{|\sigma|}{\sqrt{2}}$$

For a series of experiments:

$$\overline{\sigma_{WC}} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\sigma_i|}{\sqrt{2}}$$

This is the **experimental standard deviation of the mean**.

8 Annex B: Precipitation Filtering

Precipitation can be identified in WindCube v2 and v2.1 data via the vertical LOS wind speed. Positive speeds indicate downward motion. Some prior standards indicating data filtering due to precipitation was necessary

8.1.1 Flat Terrain

Hydrometeors advect horizontally at the same speed as the average wind. The hydrometeors experience non-linear horizontal drag forces during turbulent gusts, with characteristic decay times, after which the droplets' speed is again unified with the fluids average speed. The effect of this drag force is a function of the Stokes number of the particles. It is well known that hydrometeors generate lidar backscatter. In the WindCube, this backscatter is evident in the vertical wind speed measurements, showing positive velocities for falling rain or snow.

The effect of drag forces on horizontal wind speeds during gusts is quite small, and empirical data in Section 5 show that horizontal wind speed measurements reconstructed from hydrometeor backscatter, or a mixture of hydrometeor and aerosol backscatter show are well-correlated to collocated met masts. Note that cup anemometers also have slight, precipitation-induced low biases which are described in the literature. In summary, it is not necessary to filter the data based on vertical wind speed or wiper count.

8.1.2 Complex Terrain

In complex terrain it is possible to have large vertical wind speeds caused by terrain, and not by hydrometeors. In this case, it is recommended that you examine the distribution of vertical wind speeds in each wind direction sector to understand the possible inflow angles at the site. Again, precipitation filtering is not advised.

8.1.3 Precipitation Filter (experimental)

In some circumstances, it may be required to filter precipitation **from the vertical wind speed**, or to identify times of precipitation. For example, computations of inflow angle using the vertical wind speed should not include precipitation contamination. Vaisala's recommended practice for WindCube v2 and v2.1 filtering in this circumstance is to adjust the threshold for each 20° sector based on the distribution of vertical wind speeds in each sector using the following formula:

Vertical Wind Speed Filter Threshold:

$$\boldsymbol{w_{Threshold,Sector}} = \boldsymbol{Median}(\boldsymbol{w_{sector}}) + \mathbf{5} * \boldsymbol{Median\ Absolute\ Deviation}(\boldsymbol{w_{sector}})$$

This dynamic threshold ensures that orographically driven or convectively driven flows are not falsely flagged as precipitation, capturing the primary distribution mode of vertical wind speeds. The robustness of Median and MAD ensure this statistic is well-behaved.

Empirical evidence shows that the precipitation-influenced wind speeds are equivalently accurate to those in clear conditions.

During the WindCube v2 IEC Classification, a small sensitivity to precipitation was observed, smallest of the flagged environmental parameters. No sensitivity to precipitation was observed during the IEC Classification of WindCube v2.1.

9 Annex C: Data Sources

Our data sources for the WindCube Data Filtering Guidelines include two large measurement campaigns.

9.1 Manufacturing Tests Against Golden Lidar 2012 – 2017

The factory validation data from both the test and Golden WindCubes were sorted into the following CNR and Availability bins:

$Avail_{min}$	$Avail_{max}$	CNR_{min}	CNR_{max}	$Avail_{avg}$	CNR_{avg}	Campaigns	N_{avg}
0%	20%	-23 dB	-22 dB	7.04	-22.28	30	8.91
0%	20%	-22 dB	-17 dB	7.36	-19.61	153	6.66
0%	20%	-17 dB	-12 dB	8.05	-14.69	116	5.51
0%	20%	-12 dB	-7 dB	9.78	-9.71	33	4.60
0%	20%	-7 dB	0 dB	9.93	-3.97	4	3.50
20%	40%	-23 dB	-22 dB	29.24	-22.10	3	5.67
20%	40%	-22 dB	-17 dB	30.16	-19.61	64	4.42
20%	40%	-17 dB	-12 dB	30.05	-14.69	40	4.13
20%	40%	-12 dB	-7 dB	30.48	-9.87	32	3.63
20%	40%	-7 dB	0 dB	30.44	-4.72	11	3.64
40%	60%	-23 dB	-22 dB	47.50	-22.01	1	4.00
40%	60%	-22 dB	-17 dB	50.75	-19.94	72	4.68
40%	60%	-17 dB	-12 dB	50.07	-14.56	25	4.48
40%	60%	-12 dB	-7 dB	50.41	-9.75	43	3.88
40%	60%	-7 dB	0 dB	50.64	-4.78	16	4.44
60%	80%	-23 dB	-22 dB	--	--	0	--
60%	80%	-22 dB	-17 dB	71.18	-19.88	104	5.27
60%	80%	-17 dB	-12 dB	71.10	-14.49	39	3.63
60%	80%	-12 dB	-7 dB	70.96	-9.25	36	3.82

60%	80%	-7 dB	0 dB	70.72	-3.97	37	4.12
80%	95%	-23 dB	-22 dB	--	--	0	--
80%	95%	-22 dB	-17 dB	89.50	-19.83	170	7.06
80%	95%	-17 dB	-12 dB	89.12	-14.66	80	4.16
80%	95%	-12 dB	-7 dB	89.04	-9.49	60	4.38
80%	95%	-7 dB	0 dB	89.46	-3.58	99	5.27
95%	100%	-23 dB	-22 dB	--	--	0	--
95%	100%	-22 dB	-17 dB	99.65	-18.36	579	74.32
95%	100%	-17 dB	-12 dB	99.95	-14.32	662	202.85
95%	100%	-12 dB	-7 dB	99.97	-9.75	641	155.68
95%	100%	-7 dB	0 dB	99.91	-4.24	511	54.79

Table 6: Groupings for evaluation of precision as a function of CNR and 10-minute availability

- Sample number in each comparison (10-Minute Avail Bin, CNR Bin, Range Gate, System) ranges between n=3 and n=1619
  - Mean sample number: 127 (21 hours)
  - Median sample number: 55 (9 hours)
- The KPI Boxplots show hinges at the 25th and 75th percentiles, and whiskers at a maximum and minimum of 1.5 X Interquartile Range (IQR) beyond the hinges.

Regression Slope

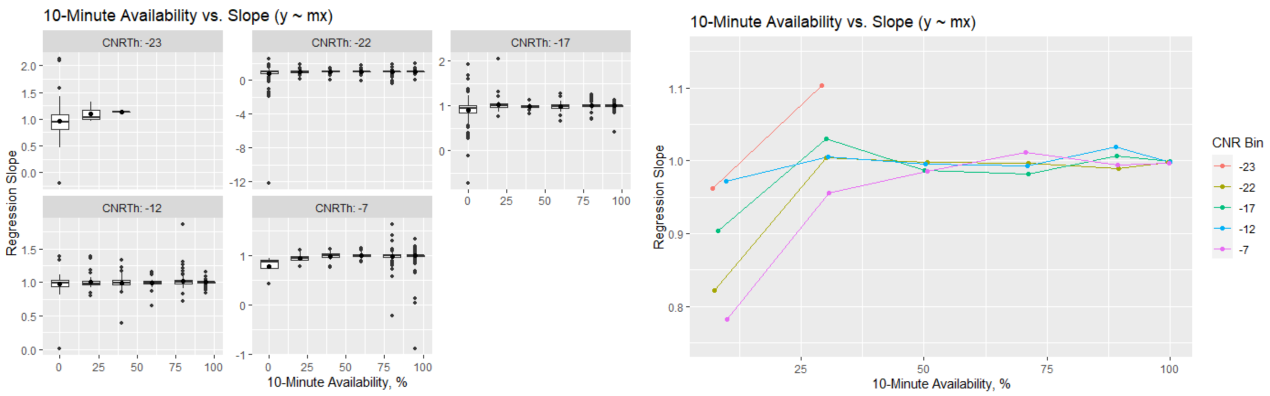

Figure 7: (A) Boxplots of least squares linear regression slopes vs. 10-Min. Avail., faceted by CNR. (B) Regression slope distribution means vs. 10-Min. Avail., grouped by CNR

Regression Intercept

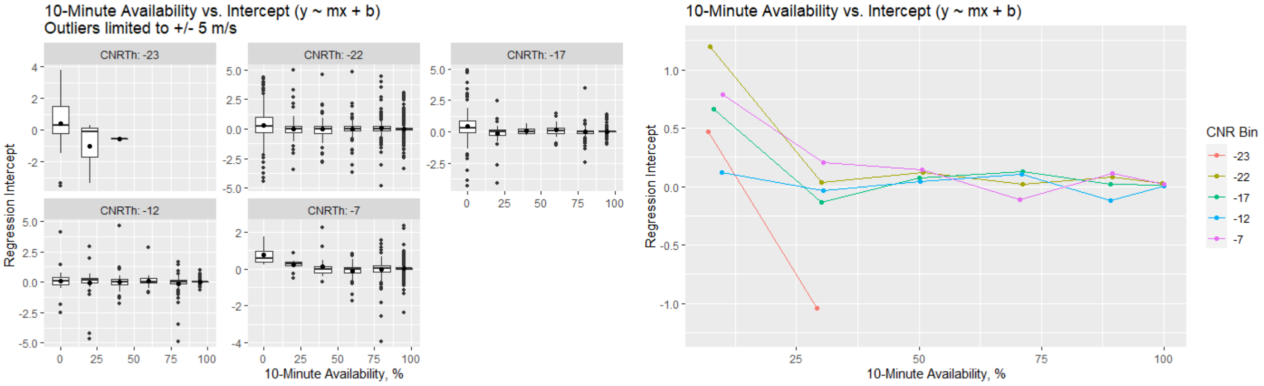

Figure 8: (A) Boxplots of least squares linear regression intercepts vs. 10-Min. Avail., faceted by CNR. (B) Regression intercept distribution means vs. 10-Min. Avail., grouped by CNR

Mean Difference

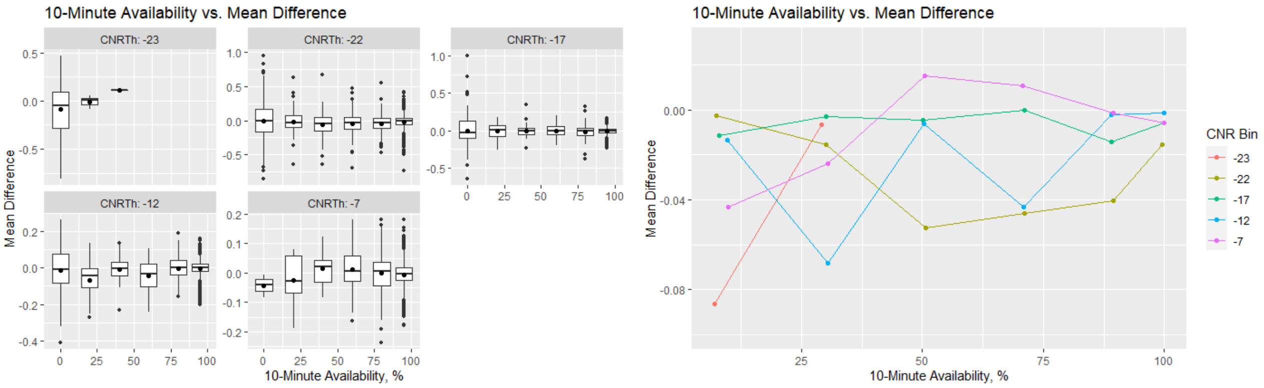

Figure 9: (A) Boxplots of mean difference vs. 10-Min. Avail., faceted by CNR. (B) Mean difference distribution means vs. 10-Min. Avail., grouped by CNR

Correlation Coefficients

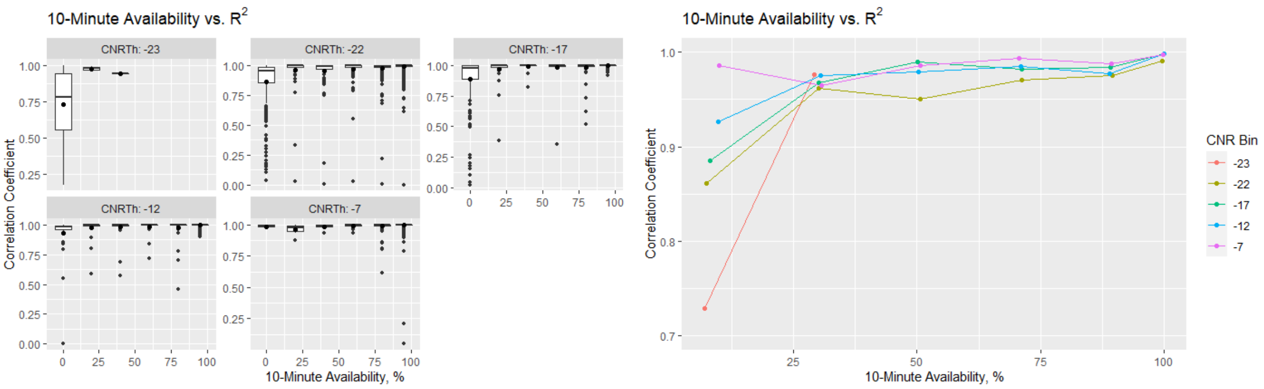

Figure 10: (A) Boxplots of R2 vs. 10-Min. Avail., faceted by CNR. (B) R2 distribution means vs. 10-Min. Avail., grouped by CNR

Experimental Standard Deviation

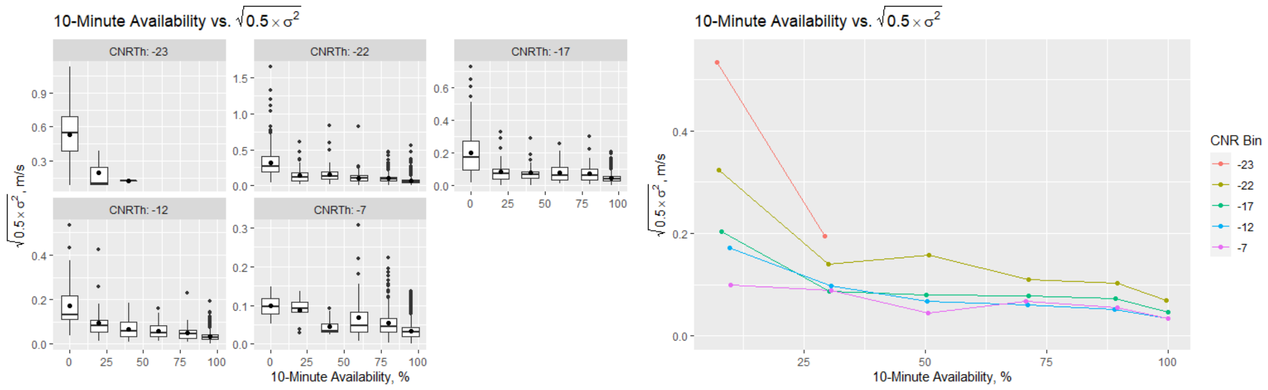

Figure 11: (A) Boxplots of Estimated Precision vs. 10-Min. Avail., faceted by CNR. (B) Estimated precision means vs. 10-Min. Avail., grouped by CNR

Sample Number and Standard Deviation of Mean Bias Error

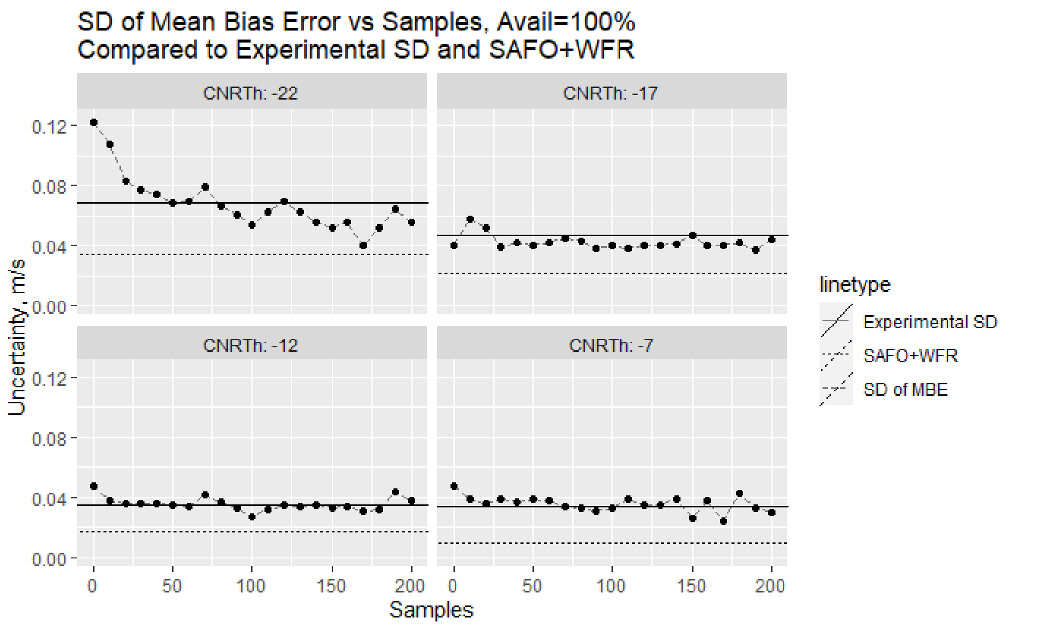

Figure 12: Standard deviation of mean bias error (Fig 9a, right-most boxes) expanded as function of sample number, compared to experimental standard deviation and SAFO+WFR, all at 100% Availability

Model Fit to (Availability, CNR) Bin Averages

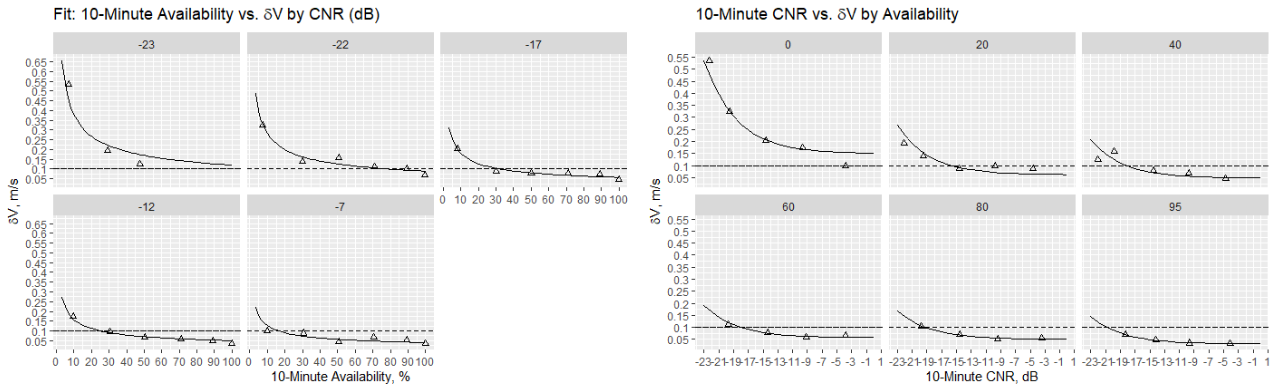

Figure 13: (A) (Avail, CNR) bin averages and model fits faceted by CNR bin (B) faceted by Availability

9.2 Precipitation: Hybrid Wind Field Reconstruction Validation 2019-2021

To estimate the accuracy of the WindCube during precipitation events, we use a dataset created in development of WindCube v2.1. Sixteen (16) WindCubes were tested at five wind measurement certification test sites. At each site, the sector-wise vertical wind speed filters described in Section 6.1.3 were derived, and the data was divided into Precipitation and Clear groups. The rain events constituted 3.1% - 8.2% of the datasets, depending on the site, and rain events were cross-checked using nearby met stations, or co-located precipitation sensors on site.

- Sample numbers for clear conditions are considerably higher than precipitation conditions.
- Those precipitation regressions with n < 36 (6 hours) are excluded from the analysis
- Filtering: 80% 10-Minute Availability, No CNR threshold.

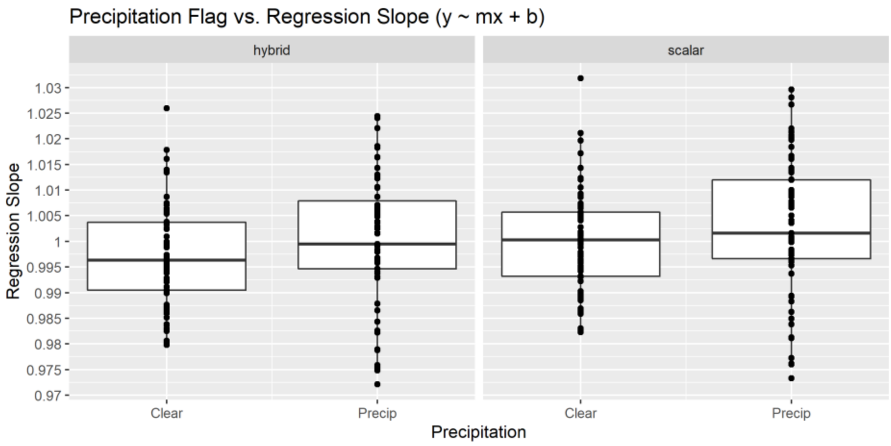

Figure 14: Boxplots of Regression Slope vs. Precipitation Flag, faceted by WFR type

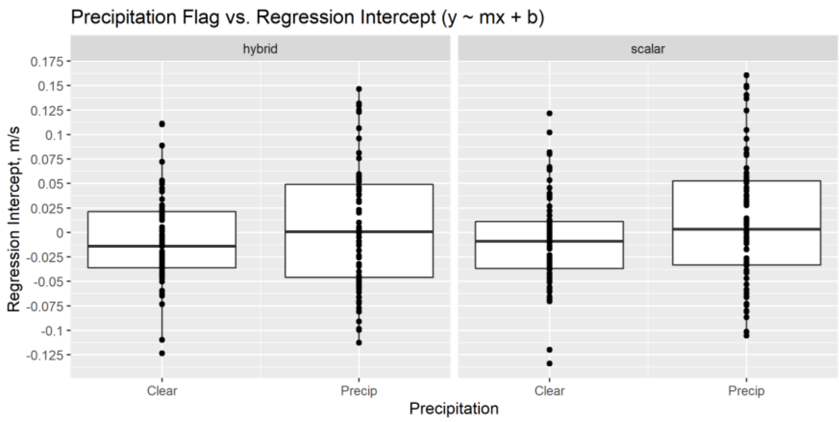

Figure 15: Figure 6: Boxplots of Regression Intercept vs. Precipitation Flag, faceted by WFR type

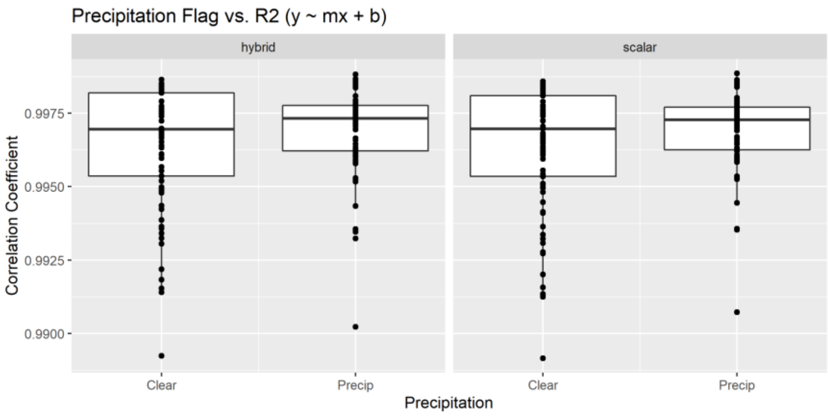

Figure 16: Boxplots of Correlation Coefficient vs. Precipitation Flag, faceted by WFR type

- Regression slope, intercept, and R2 all show excellent performance in precipitation conditions, equivalent to the KPIs for clear conditions for both hybrid (WindCube v2.1) and scalar (WindCube v2) WFR methods

9.3 IEC Classifications

To use remote sensing devices for contractual wind turbine power performance testing, devices must have their atmospheric sensitivities evaluated following the IEC 61400-12-1. This process is referred to as device Classification. Classification tests have been carried out for both the WindCube v2 and WindCube v2.1.

Copies of the full classification reports are available upon request.

The observed environmental sensitivities also mostly can be understood from lidar physical principles, particularly wind shear and turbulence intensity.

9.3.1 WindCube v2

WindCube v2's IEC Classification was carried out in 2018 by Deutsche WindGuard using three measurement sites in Northern Germany. Results showed sensitivities to wind shear, turbulence intensity, wind direction, and precipitation. WindCube v2 uses scalar wind field reconstruction.

9.3.2 WindCube v2.1

WindCube v2.1's IEC Classification was carried out in 2021 by Deutsche WindGuard using three measurement sites in Northern Germany. Results showed sensitivities to wind shear, turbulence intensity, and wind direction. WindCube v2.1 uses hybrid wind field reconstruction. Notably, the environmental sensitivity to turbulence intensity was reduced by an order of magnitude, in close agreement with hybrid wind field reconstruction theory.

10 References

Cariou, J.-P., Allain, P. E., and Thobois, L., “Wind speed precision estimation with coherent pulsed doppler lidars,” The 29th International Laser Radar Conference (ILRC29) (2019).

Clive, P. J. M. "Compensation of vector and volume averaging bias in lidar wind speed measurements." IOP Conference Series: Earth and Environmental Science. Vol. 1. No. 1. IOP Publishing, 2008.

Dentler, FU. The effect of rainfall on measurements of mean wind speed with cup anemometers in the surface layer at sea. Boundary-Layer Meteorol 14, 123–130 (1978). https://doi.org/10.1007/BF00123993

Franke, K. “Summary of Classification of Remote Sensing Device. Type: Leosphere Windcube V2.1” Technical Report. Varel, Germany (2020)

Franke, K. “Summary of Classification of Remote Sensing Device. Type: Leosphere Windcube” Technical Report. Varel, Germany (2018)

IEC-TC88 MT12-1, Standard IEC 61400-12-1:2017, *Wind Energy Generation Systems-Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines* (International Electrotechnical Commission, Geneva, CH, 2017).

Mazoyer, P., Hermaszewski, C., and F. Rebeyrat. “Assessment of the repeatability and stability of Windcube performances.” Resource Assessment 2017. Edinburgh: WindEurope, 2017

Rosenbusch, Peter, et al. "Wind speed reconstruction from mono-static wind lidar eliminating the effect of turbulence." Journal of Renewable and Sustainable Energy 13.6 (2021): 063301.

Rye, B. J. and R. M. Hardesty, "Discrete spectral peak estimation in incoherent backscatter heterodyne lidar. I. Spectral accumulation and the Cramer-Rao lower bound," in IEEE Transactions on Geoscience and Remote Sensing, vol. 31, no. 1, pp. 16-27, Jan. 1993, doi: 10.1109/36.210440.

Banakh, Viktor, and Igor Smalikho. Coherent Doppler wind lidars in a turbulent atmosphere. Artech House, 2013.

Evaluation of Site-Specific Wind Conditions. Version 2. MEASNET (2016)