Windspeed recording process and related issues

Vincenzo Ferrazzano *

April 30, 2010

Abstract

Recording process for windspeed data and subsequent data manipulations are discussed. Here we show why those manipulations should be taken into account in modeling. Those operation can lead to artifacts that can be misinterpreted or affect the estimation of physically meaningful quantities.

1 Experimental facts and operative hypothesis

We shall recap here some well known stylized facts about turbulence. Those are observed in almost every fully developed turbulent dataset.

(i) Two-thirds law. In a turbulent flow at very high Reynolds number, the mean square velocity increment $\mathbb{E}\left[v(x+l)-v(l)\right]^2$ between two points separated by a distance l behaves approximately as the $2/3$ power of the distance. That is equivalent to saying that the velocity power spectral density must decay exponentially with exponent $-5/3$. This happens on an interval length scales over which energy is transferred and dissipation due to molecular viscosity is negligible.

(ii) Law of finite energy dissipation. If, in an experiment on turbulent flow, all the control parameters are kept the same except for the viscosity, which is lowered as much as possible, the energy dissipation per unit mass behaves in a way consistent with a finite positive limit.

Those experimental facts are explained in greater detail in Chapter 5 of Frisch (1996); another source is McDonough (2007). Fact (ii) is almost always observed, but an eventual contradiction of this fact would not lead to a violation of any known law of nature. All the physical theory on turbulence is developed in a spatial context, i.e. dependence on time is not considered. But most experimental data are drawn from a fixed sensor, and measured as a time series. Then it is usual to employ the (questionable) Taylor hypothesis:

(iii) Taylor "frozen field" hypothesis. Assume that the velocity field can be decomposed in a mean flow U and a time varying part v':

$$v(x,t) = v'(t, x-Ut) + U.$$

Then, if $\sqrt{\mathbb{E}\left[v'^2\right]}/U << 1$, it is reasonable to exchange the time increment τ to the space increment l using the relation $l = U\tau$.

*Center of Mathematical Sciences, Lehrstuhl f§r Mathematische Statistik, Parkring 13, 85748 Garching, Germany, email: ferrazzano@ma.tum.de, http://www-m4.ma.tum.de/pers/ferrazzano/

For our purposes we can use this law the other way around: we are interested in the fixed-point, time-wise modelling, and we will use the frozen field hypothesis to bring the physical, space-domain living laws, into the time domain.

(iv) Scaling of increments. Let l_0 be a length typical of the device generating the turbulence, then for l in the order of l_0 the probability density function (p.d.f.) of increments is approximately Gaussian, while on smaller scale the p.d.f. exhibits exponential tails. They become even heavier when the scale decrease.

For our data we point out that in atmospheric turbulence the integral scale will be roughly in the order of 50 m, i.e. we are almost always in the heavy tailed case (see Figure 1). Barndorff-Nielsen and Schmiegel showed that the data fit well the normal inverse Gaussian (NIG) distribution with a pronounced asymmetry in the tails.

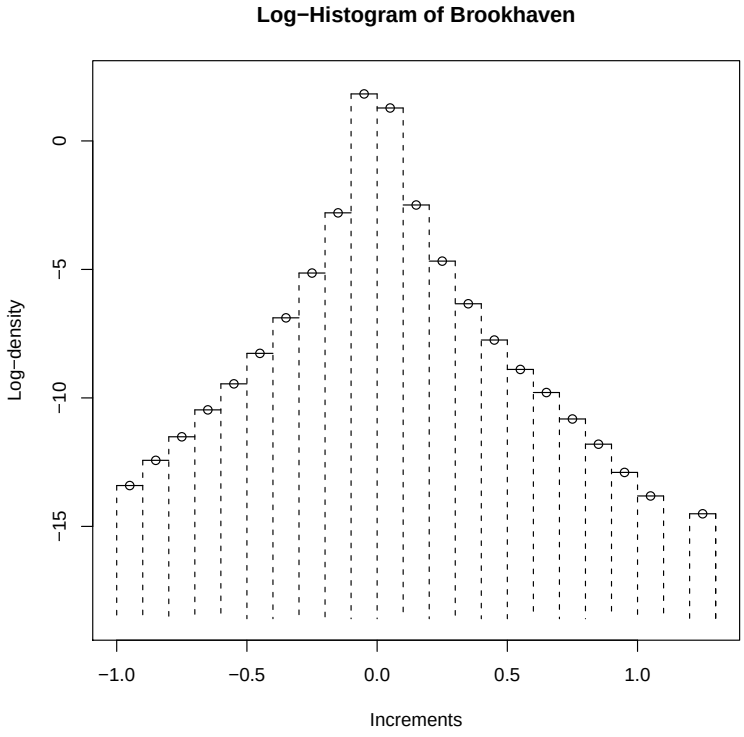

Figure 1: Log-histogram of the increments of the Brookhaven dataset. The histogram is logarithmic in order to emphasize the heavy tails.

2 Experimental data

Our data of atmospheric turbulence come from the Brookhaven National Laboratory, located in Upton, New York on Long Island. As usual in the field of meteorology, we consider only the mean flow component of the windspeed. It means that, given a sample of a three dimensional

flow $\mathbb{R}^3 \ni u(t) = (u_x, u_y, u_z)$ we calculate the mean flow

$$\bar{u} = \frac{1}{T} \sum_{i=0}^{N} u(t_i) \quad 0 = t_0 < t_1 < \ldots < t_N = T.$$

Then the mean flow component u_t is

$$v(t) = u(t) \cdot \frac{\bar{u}}{||\bar{u}||}$$

where $\cdot$ denotes the usual inner product in $\mathbb{R}^3$ and $||\cdot||$ is the Euclidean norm. Normally, wind conditions are considered to be stable (read. stationary) up to 15 minutes, and IEC standards impose that all the averaging procedure involving time are performed on a time reference of 10 minutes.

2.1 Brookhaven dataset

The dataset considered is the Brookhaven dataset: it is a high frequency dataset, recorded by Drhuva employing a hot wire anemometer; a full account of the data can be found in Drhuva (2000). Brookhaven dataset consists of $20 \cdot 10^6$ data points sampled at 5kHz (i.e. $5 \cdot 10^3$ points per second). The acquisition process with this kind of technique will be the object of a lengthy discussion in the next section. The dataset was kindly provided by Ole Barndorff-Nielsen and Jürgen Schmiegel from Aarhus University (Denmark).

3 Noise, errors and manipulation

It is quite common in physics and engineering to consider the omnipresent instrumental error to be Gaussian-distributed, when many sources of errors of comparable entities are concurring. Main sources of error for our kind of data are:

- High frequency errors: errors caused by the change in hot wire behavior at high frequency.
- Spatial resolution errors: errors caused by spatial averaging of the flow field.
- Disturbance errors: errors caused by the probe interfering with the flow field.
- Calibration measurement errors: errors in measuring the calibration flow parameters and hot wire voltages.
- Calibration equation errors: errors due to the fitting of a calibration equation, as well as solution of the calibration equation.
- Calibration drift errors: errors caused by variations in calibration over time and over on/off switching of the feedback circuitry, as well as by probe contamination.
- Approximation errors: errors caused by assumptions about the flow field that are used to solve the calibration equations.
- Directional error: the geometry of the probe does not allow measurements to be acquired with same accuracy in all directions.

The last issue can be very important in atmospheric anemometry, since the wind direction is swinging and turning. All the consulted sources agree on the fact that all measurements are to be considered affected by errors, which can be kept at bay, but not eliminated, by a correct preparation of the experiment. For further reference check Tropea (2008) and Bruun (1995).

3.1 Data acquisition process: hotwire anemometer

Here we schematize and describe operations performed to collect windspeed data. The actual output of the hotwire anemometer is the voltage $V^0(t),\ t \in [0, T]$. We will denote in this section with V a voltage and with v a windspeed. To stress, when a quantity X is a continuous time signal, we will use $X(t)$. For the discrete time values, the notation X_i is used.

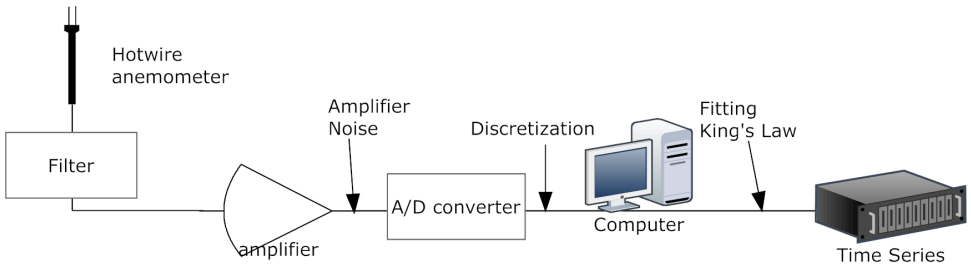

Figure 2: Succession of typical operation performed on the data, as reported in Bruun (1995).

As mentioned before, the output of a hotwire is not a windspeed, but the voltage $V^0(t)$ between two ends of the anemometer. This signal will be filtered by applying a low-pass filter, with cut off frequency set at half (or less) of the sampling frequency. This action is based on the Shannon-Nyquist theorem that assure that a good reconstruction of the signal in the time domain can be obtained only for information contained in the part of the spectrum with frequency $f < f_0/2$. All those operations are applied to the analog (i.e. continuous time) signal through electronic devices.

After those filtering operation we have a signal $\tilde{V}^0(t)$, which is amplified through a voltage amplifier, netting, in line of principle, an output $V^1(t) = a\tilde{V}^0(t)$, where $a \in \mathbb{R}^+$ is the so-called gain factor.

In reality the behavior of an actual amplifier is slightly more complex; two factors should always be kept in mind:

Non-linearity The relationship $V^1 = a\tilde{V}^0$ holds in the *ideal* case of a linear amplifier. Unfortunately, amplifiers are intrinsic non-linear devices, since the gain factor is not a constant, but it tends to amplify some portion of the spectrum more than others. Anyway, it is possible to build amplifiers with linear behavior in the band of interest, then this factor normally is not a concern.

Noise The amplifier generates, beside the amplified voltage, a noise, which depends on the quality of the components composing the amplifier, and also environmental conditions (mainly thermal noise).

Then the actual output of the amplifier is: $\tilde{V}^1(t) = a\tilde{V}^0(t) + n_{amp}(t)$.

After these manipulation the signals are passed through an analog-to-digital converter (ADC), in order to get a discrete sequence of values, to be used by a computer. The resolution of an ADC is expressed in bits, indicated with an integer p. Let us assume that the signal to be sampled is included in the interval $E = [V_l, V_h]$, then we call $|E| = V_h - V_l$ the voltage scale. The voltage scale can be adjusted, accordingly to the amplitude of the incoming signal, but after that it remains fixed during the experiment. Since only signals with values in E can be properly sampled, it is necessary to take E large enough in order to have

$$V_l \leq a\tilde{V}^0(t) + n_{amp}(t) \leq V_h \quad t \in [0, T],$$

otherwise the experiment has to be run again, since the exceeding values will be chopped off. Then the digital data will have a truncation level of

$$\alpha = \frac{|E|}{2^p - 1}. \tag{1}$$

Now we obtain discrete-time data V_i, for $i = 0, \ldots, [Tn]$:

$$V_i = \alpha \left\lfloor \frac{\tilde{V}^1(i/n)}{\alpha} \right\rfloor = \alpha \left\lfloor \frac{a\tilde{V}^0(i/n) + n_{amp}(i/n)}{\alpha} \right\rfloor = \alpha \left\lfloor \frac{\tilde{V}^0(i/n)}{\alpha'} + \frac{n_{amp}(i/n)}{\alpha} \right\rfloor; \tag{2}$$

where $\alpha' = |E|/a(2^p - 1)$ is the equivalent truncation level of the signal $V^0(t)$.

In general, the ADC is a non-linear device. The voltage scale $|E|$ and the amplifier gain has to be chosen in order to minimize the truncation level, but keeping an eye on the specifics of the electronic equipment, because every electronic device has precise operative ranges, where the behavior of each instrument is, to a satisfying level, linear. Usual values in the field of hotwire anemometry are $p = 12$ bit, $E = 20$ volt, with an a chosen accordingly. That means $\alpha \approx 0.0012$, i.e. the voltage increment has no more than 3 digits.

The last step to get velocity data is the calibration curve: there exists a deterministic relationship between voltage and windspeed, which has to be determined for each anemometer. An anemometer is normally tested in turbulence facility, like a wind tunnel, with a stable, constant and known speed and turbulence; then the output voltage is recorded. The procedure is repeated for some windspeeds (usually 10-30 values) and then results are interpolated, to get a calibration curve; which is, mathematically speaking, a continuous function $f_\theta : [V_l, V_h] \to [0, v_{max}]$, which depends on a vector θ of parameters.

Example 1. The most common interpolating function is given by King's Law

$$V^2 = k_0 + k_1 v^n, \quad V \in E, \tag{3}$$

where the exponent n depends on the geometry of the sensor. Then $f_\theta(V) = \left(\left(V^2 - k_0\right)/k_1\right)^{1/n}$ and the vector is $\theta = (k_0, k_1, n)$. We note that f_θ is increasing and concave. If the voltage signal had symmetric increments, the windspeed will not.

Normally this procedure is too expensive to be performed for each commercially available anemometer, and it is reserved only to those to be employed in high precision scientific measurement. For all other purposes the producer provides a standard calibration curve.

Although the discretization procedure in the voltage data is clear, in wind data it is not, since now the discretization step depends on the derivatives of the calibration curve, which normally is not given along with the time series. Then, if we plot histograms like the one in Figure 1, we are not able to see any discretization, which becomes visible only at a closer glance, like in Figure 3. Fortunately, it is possible to recover some information from the windspeed data.

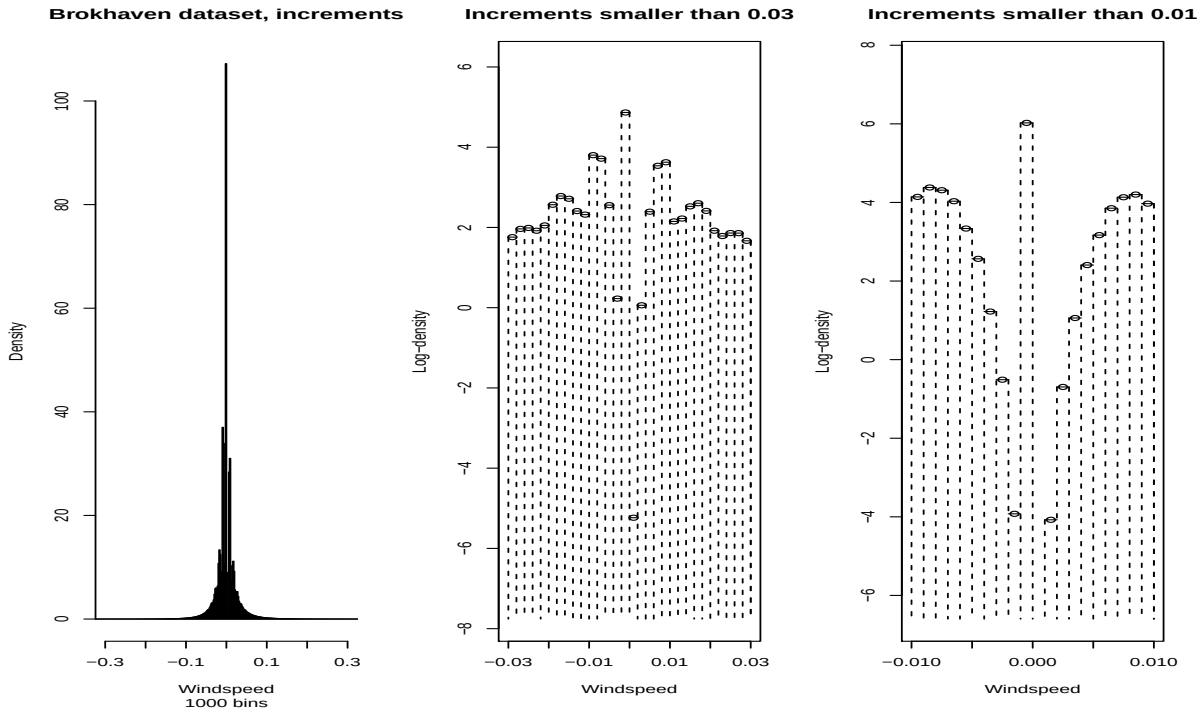

Figure 3: Closer look to windspeed increments, for the Brookhaven dataset. It turns out that 21.43% of the total increments are 0.

4 Recovering the parameters

Let us suppose now that the interpolation function f_θ is smooth enough to perform all the needed derivatives. Using Taylor's formula on $f_\theta(V)$ we get for appropriate $n \in \mathbb{N}$:

$$\Delta v = f_\theta(V + \Delta V) - f_\theta(V) = \sum_{i=1}^{K} \frac{1}{i!} f_\theta^{(i)}(V)(\Delta V)^i + o\left((\Delta V)^K\right). \tag{4}$$

we conclude from the rhs of (4) that the new discretization steps depend on derivatives of the calibration curve f at voltage V and on the discrete increments ΔV. Since Taylor's formula is precise only for small ΔV and, by smoothness of f, for small Δv, it makes sense to consider only the increments up to a threshold β. Then for all $j_i \in \mathbb{N}$ such that $|\Delta_{j_i} v| < \beta$:

$$\Delta^\beta v = (\Delta_{j_0} v, \ldots, \Delta_{j_n} v)$$

where $\Delta_j v = v_{j+1} - v_j$. Then we summarize the associated levels v^β as

$$v^\beta = (v_{j_0}, \ldots, v_{j_n}).$$

Plotting $\Delta^\beta v$ versus v^β one has a striking visual effect of the discretization effect, as depicted in Figure 4. Of course, null increments are distributed along all the measurement levels, but other increments move along solid curves, as the speed varies; this has to be attributed to the variation of the derivatives of f_θ. Then we fix V and we get discrete increments $\delta^i v := f_\theta(V + i\Delta V) - f_\theta(V)$;

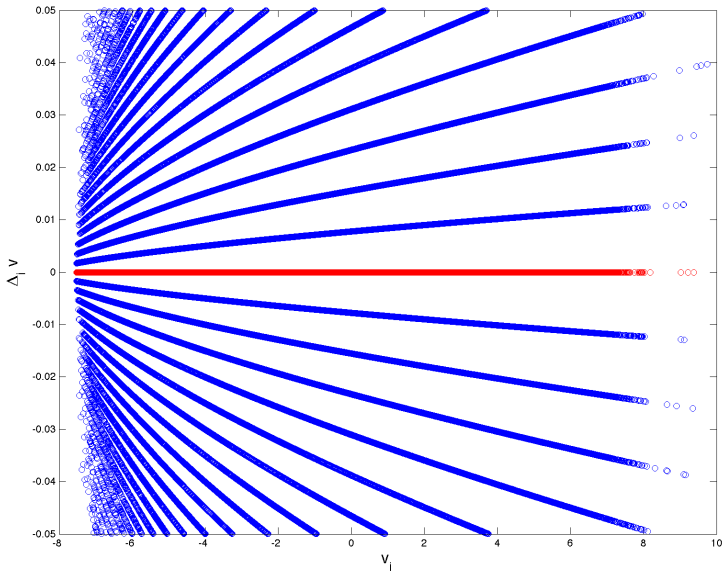

Figure 4: $\Delta^\beta v$ vs. v^β for the Brookhaven dataset, $\beta = 0.05$. The non-linear behavior of the curves is given by derivatives of King's law. Note that 92.36% of all increments of Brookhaven dataset belongs to $\Delta^{0.05}v$, including the already mentioned 23% of null increments (in red).

using Taylor's formula:

$$\begin{aligned} \delta^1 v &= \sum_{i=1}^{K} \tfrac{1}{i!} f_\theta^{(i)}(V)(\Delta V)^i + o\left((\Delta V)^K\right) \\ &\vdots \qquad\qquad \vdots \\ \delta^K v &= \sum_{i=1}^{K} \tfrac{1}{i!} f_\theta^{(i)}(V)(K\Delta V)^i + o\left((\Delta V)^K\right) \end{aligned} \tag{5}$$

Since our data are given in velocities, we perform a change of variables: $V = f_\theta^{-1}(v)$. Then, in the light of (5) we can interpret Figure 4 in two different manners:

- We fix v and we look at the different branches of the increments: here we see how the velocity can jump discretely only to some predefined values, that is, moving with step ΔV.
- We look only at the first, upper branch (i.e. the one composed by the smallest non-null increments). If we move along this curve, we have fixed ΔV and we vary v, that is, we see how the derivatives of f_θ influence the increments.

The first outlook gives us an immediate feeling of the discreteness of our data, but the second one is more productive. Since $\Delta V \sim (2^{12} - 1)^{-1} = 2.4420 \cdot 10^{-4}$ we can hope to reach a good approximation with low order Taylor series.

We propose the following strategy to recover the parameter vector θ.

Take increments $\Delta^{(1)}v := \{\Delta v : \text{belongs to the first branch}\}$ and the levels $v^{(1)} := \{v : \text{belongs to the first branch}\}$. Furthermore, let $\phi = (\theta, \Delta V)$ be the extended parameter vector of the Taylor series and let $\phi^{(0)}$ be a guess on ϕ, which we take as initial values.

Then we perform a non-linear least squares fitting on the first order Taylor series

$$\Delta^{(1)}v = f'_\theta(f_\theta^{-1}(v^{(1)}))\Delta V,$$

returning the estimate $\phi^{(1)}$. If the estimate is too crude, we may need to take higher order Taylor polynomials into account, getting the estimate, for instance, $\phi^{(2)}$ from a least squares fitting on the second order:

$$\Delta^{(1)}v = f'_\theta(f_\theta^{-1}(v^{(1)}))\Delta V + f''_\theta(f_\theta^{-1}(v^{(1)}))\frac{\Delta V^2}{2}$$

improving $\phi^{(1)}$, and so on. If the function f_θ is complicated, it may be a good idea to proceed step-by-step, but if f_θ is polynomial, we can use directly the Taylor series of order equal to the order of f_θ.

King's law and Brookhaven data We will use this strategy on the Brookhaven dataset, in order to recover its true discretization pattern. As stated in Drhuva (2000), the Brookhaven data were obtained using King's law. We report here the derivatives up to order 3:

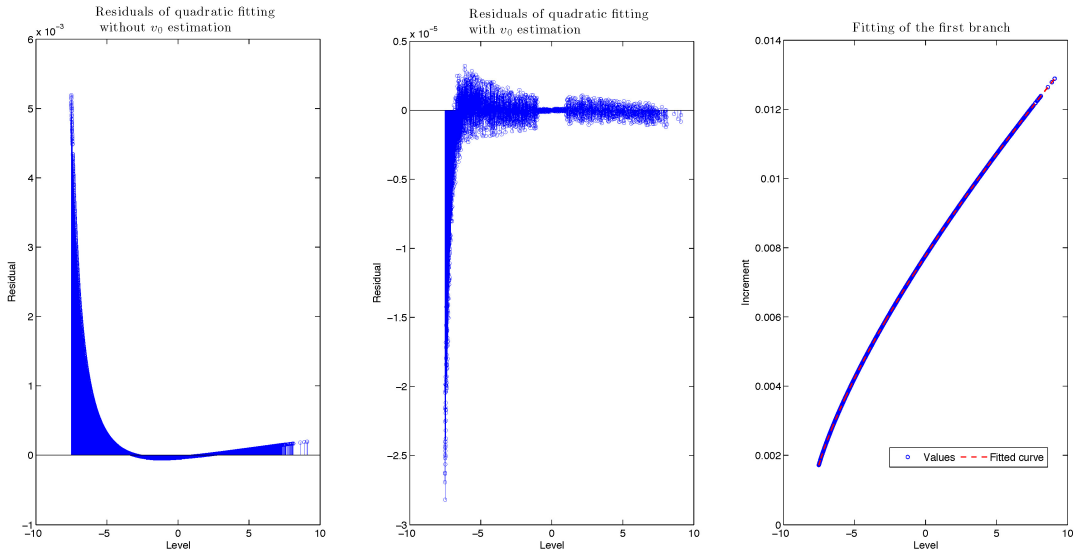

Figure 5: Figures relative to the recovery of the parameter vector θ for the Brookhaven dataset, using second order fitting.

$$f'(f^{-1}(v)) = \frac{2v^{1-n}\sqrt{k_0 + k_1 v^n}}{k_1 n}$$

$$f''(f^{-1}(v)) = \frac{2v^{1-n}}{k_1 n} + \frac{4\left(\frac{1}{n} - 1\right)v^{1-2n}\left(k_0 + k_1 v^n\right)}{k_1^2 n}$$

$$f'''(f^{-1}(v)) = \frac{12\left(\frac{1}{n}-1\right)v^{1-2n}\sqrt{k_0+k_1v^n}}{k_1^2 n} + \frac{8\left(\frac{1}{n}-2\right)\left(\frac{1}{n}-1\right)v^{1-3n}\left(k_0+k_1v^n\right)^{3/2}}{k_1^3 n}$$

For the first branch the behavior should be given mainly by the first derivative, which allows us to make a guess on the parameter n, which gives the shape of the curve. For small values of v the curve goes to zero, and for big values of v it grows somehow linearly, then we can suppose $n \in (0,1)$, with an initial guess $n^{(0)} = 0.5$.

Since the mean value of the windspeed was unfortunately subtracted from the Brookhaven data, we have to consider, instead of the actual velocities v, a Reynolds-like decomposition

$$v(t) = u(t) + v_0, \quad t \in [0,T],$$

where v_0 is the unknown average windspeed. Since the derivatives depend in a nonlinear way on the values of the windspeed, this information is important and it has to be recovered.

In Drhuva (2000), Table 1, for entry 3 (BNL) $v_0 = 8.3$ is reported, but it was calculated on 4×10^7 values, while our dataset has "only" 2×10^7 datapoints. Two strategies are viable: taking the average speed in Drhuva (2000) for real, or use it as a guess, increasing the numbers of parameters to be fitted by 1.

We tried both, getting similar parameters; it is very interesting to notice how a slight variation in the value of v_0 causes a relatively better fit, accordingly to the residuals listed in Table 1. However, the really close estimation of average windspeed on only of the half dataset is, at least, a good signal of the stationarity of this dataset.

Table 1: Non-linear least squares fitting for Brookhaven dataset.

	k_0	k_1	n	ΔV	v_0	max(res)	max(res/$\Delta^{(1)}v$)
$\phi^{(0)}$	12	12	0.5	$10 \cdot (2^{12}-1)^{-1}$	8.3	-	-
$\phi^{(1)}$	13.0084	10.4190	0.5005	$13,40 \times 10^{-4}$	8.3	8.78×10^{-6}	0.0051
$\phi^{(2)}$	12.9990	10.4449	0.5010	$13,40 \times 10^{-4}$	8.3	8.97×10^{-6}	0.0052
$\phi^{(1)}$	13.0386	10.5953	0.5000	13.31×10^{-4}	8.3115	9.87×10^{-9}	3.23×10^{-6}
$\phi^{(2)}$	13.0316	10.5653	0.4999	$13,32 \times 10^{-4}$	8.3120	1.05×10^{-8}	3.21×10^{-6}

Once we have recovered θ, we can proceed to get the voltage data, and compare them to the windspeed data. In Figure 6 it is rather obvious that the discretization in the voltage data is generating the unexpected behavior in the center of the histogram of the Brookhaven data.

From the voltage data we obtain that $\Delta V = 0.0011$, which is not that far from the one estimated by the least squares fitting. We do stress that the estimate ΔV do not enter in King's law, but it was needed, since it gives the scaling in the Taylor series. That is translated in $E = 0.0011 \cdot (2^{12}-1) \approx 4.5V$, which is in agreement with the voltage data since $\max_i(V_i) - \min_i(V_i) = 2.8763$. That indicates that experimenters took a large, but reasonable voltage scale, to be sure to be able to sample the whole signal, sacrificing some accuracy. Another nice thing is that the voltage histogram exhibits symmetric tails.

4.1 Comparison between voltage and windspeed

Now we have voltage and windspeed, we can perform the standard analysis on both of them, in order to understand how we can profit of the addiction knowledge of the sampling process. It

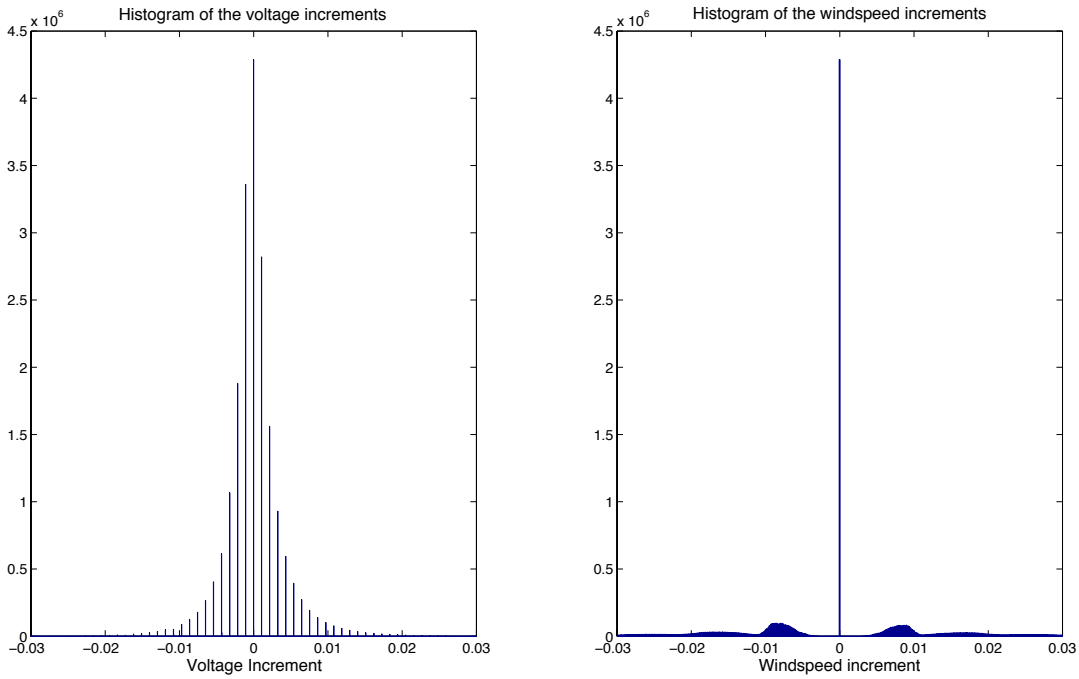

Figure 6: Histogram of increments for Brookhaven data, for windspeed and voltage (20000 bins). In the voltage histogram the discretization structure of the data is evident.

comes for free that the voltage (and therefore the windspeed) has to enjoy a rather strong form of continuity, since a circuity fed with a discontinuous input will react anyway with a smoother output. It comes as well from the physical limitations of the ADC converter that the voltage has to take value in a compact of limited measure.

We fitted a NIG distribution to the voltage increments data, as suggested by Barndorff-Nielsen and Schmiegel. It turns out that the data can be well fitted using only two parameters α (which commands the tail heaviness) and δ (scaling factor), since the localization μ and the skewness β are rather small. Especially the skewness β has to be considered in comparison to α, thought the quantity $\gamma = \sqrt{\alpha^2 - \beta^2}$. It turns out from Table 2 that $\alpha/\gamma \approx 1.002$, then we can pay no attention to the skewness parameter.

Table 2: Parameter of the NIG distribution for the increments of Brookhaven voltage data.

α	β	μ	δ	γ
0.481912	0.0442513	-0.00014419	0.00147126	0.4799

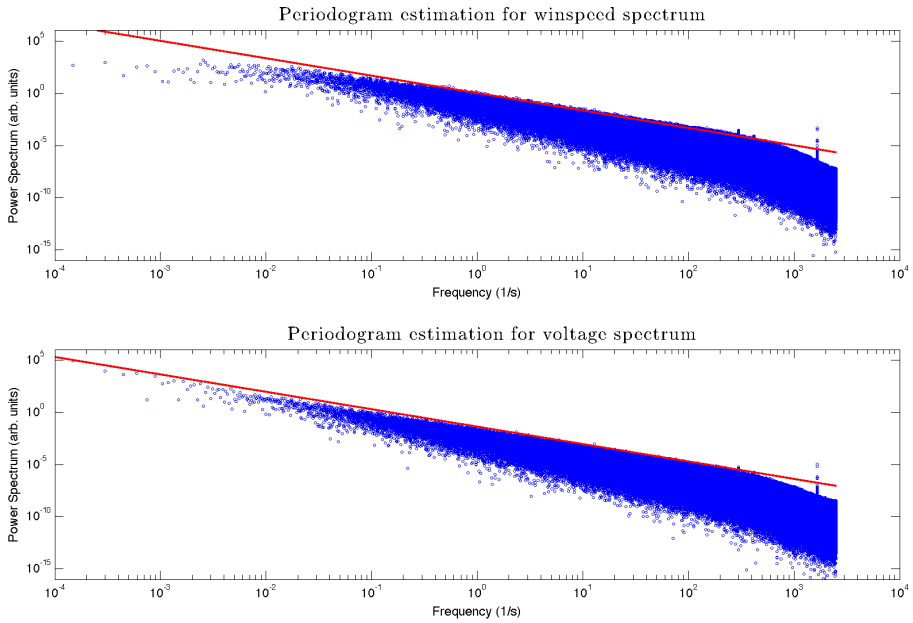

Figure 7: Energy spectrum comparison for voltage and windspeed. The red line has -5/3 slope predicted by Kolmogorov (1941) . We can see how the voltage has the 5/3 slope even for the smaller decades. In the latter part we can see a change of slope, rather than a nonlinear behavior.

5 Filters

For a $g \in L^2$ we denote by

$$F(g)(\omega) := \int_{-\infty}^{\infty} g(x) e^{i\omega x} dx, \quad \omega \in \mathbb{R},$$

the Fourier transform of g, and by $F^{-1}(\cdot)$ its inverse.

The natural measure in the frequency domain is the frequency, which has dimension $[s^{-1}]$; it is customary to shorten the notation working with the angular frequency $\omega = 2\pi f$, which is measured in $[rad\ s^{-1}]$.

Definition 5.1 (Ideal low-pass filter)**.** An ideal low-pass filter is a positive deterministic function $h \in L^2$, such that its Fourier transform is

$$F(h)(\omega) = 1_{\{\omega \leq \omega_c\}}$$

where ω_c is the cutoff frequency. $P := \{\omega \leq \omega_c\}$ is the pass-band, and $S := \mathbb{R}^+ \backslash P$ is the stop band.

Then the idea is to use such filter to remove all the spectral components present at frequencies higher than ω_c. Such ideal filters have some drawbacks, nor are practically attainable, so engineers develop circuits, which have a similar effect, approximating the ideal filter.

Definition 5.2 (Non-ideal filter)**.** A non-ideal low-pass filter is a positive deterministic function $h \in L^1$, such that its Fourier transform satisfies:

- $\lim_{\omega \to 0} F(h) = 1$
- $\lim_{\omega \to \infty} F(h) = 0$

The value of 1/2 in the definition of cutoff is arbitrary and it comes from engineering practice. That means the power spectrum of the filtered data at ω_c is reduced by a factor 1/2 with respect to the original one.

Example 2. The simplest low pass filter is the Butterworth filter:

$$|h(\omega)|^2 = \left(1 + \left(\frac{\omega}{\omega_c}\right)^{2n}\right)^{-1}, \quad \omega \in \mathbb{R},$$

where n is the order of the filter, and ω_c is the cutoff. The Butterworth filter has two main features:

$$\text{Maximal flatness:} \quad \lim_{\omega \to 0} \frac{d^k}{d\omega^k} |h(\omega)| = 0, \ k < 2n - 1.$$
$$\text{Linear roll-off:} \quad \lim_{\omega \to +\infty} \frac{\log |h(\omega)|}{\log \omega} = -n$$

The first feature is an ideal one, because it ensures that the passband have no ripple or distortion. The linear roll-off in the stop band is the mayor drawback in this filter. If a steeper roll-off is needed, other more elaborate filters must be employed; however Butterworth filter is widely employed in the engineering practice, when a not surgical cutoff is needed. We refer to Bianchi and Sorrentino (2007) for further details.

Drhuva (2000) reports that he applied two times a Butterworth filter to his data, with cutoff frequency at most equal to the half of the sampling frequency, without reporting precisely the order or the cutoff frequency for the dataset we analyzed.

Regardless of this information, we can say something about the effect of the filter. If the filter had a cutoff exactly at 2500Hz, the spectrum at this frequency is reduced by a factor 1/4; contributing to the bending that can be seen in Figure 7. The effect of this bending on the data has to be investigated.

6 Conclusion and to-dos

We have found a way to move from windspeed data from the original voltage data, which show in a clear way the discretization effect of the sampling. Since we are only one, deterministic, step away from the windspeed data, it suggests that can be useful to study directly the voltage data, since the discretization issue is clear in the voltage time series, while is not in the "raw" windspeed. For a given velocity model, the presence of artifact induced by

- filtering
- noise
- discretizing

can affect the estimation of parameters from the data. The latter two have been already considered in the world of financial mathematics, like in Delattre and Jacod (1997) and Zhang et al. (2005), and that pose the same questions in a different context. On the other hand, the question on filtering is completely new, since is something typical of the engineering world and it is normally considered (in that context) as an operation without any drawback. But as already noted in the literature on spectral analysis (e.g. Priestley (1981), pg. 516) there is a duality on spectrum and autocorrelation. The filtering operation modifying the spectrum making it "shorter" (i.e. making the low frequency part more significative than the high one), adding some spurious long range dependency on the data.

7 Acknowledgement

I would like to thank Roman Reß for showing me how the sampling process works, giving me some confirmation about physical concepts and intuitions. Furthermore I would like to thank Ole Barndorff-Nielsen and Jürgen Schmiegel for the Brookhaven dataset and for all the discussions Moreover I am thankful for fruitful discussions with my PhD supervisor Claudia Klüppelberg.

References

Barndorff-Nielsen, O. E. and Schmiegel, J.: 2008, Time change, volatility and turbulence, *Technical report*, Thiele Center for Applied Mathematics in Natural Science, Aarhus.

Bianchi, G. and Sorrentino, R.: 2007, *Electronic filter: simulation and design*, McGraw-Hill, New York.

Bruun, H.: 1995, *Hot-Wire Anemometry: Principles and Signal Analysis*, Oxford University Press, Oxford.

Delattre, S. and Jacod, J.: 1997, A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors, *Bernoulli* **3**(1), 1–28.

Drhuva, B. R.: 2000, *An experimental study of high Reynolds number turbulence in the atmosphere*, PhD thesis, Yale University.

Frisch, U.: 1996, *Turbulence: The Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge.

Kolmogorov, A. N.: 1941, The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers, *Proceedings of the USSR Academy of Sciences* **30**, 229–303.

McDonough, J. M.: 2007, Introductory lectures on turbulence. `http://www.engr.uky.edu/~acfd/lctr-notes634.pdf`.

Priestley, M. B.: 1981, *Spectral Analysis and Time Series*, Vol. 1, Academic Press, London.

Tropea, C., Yarin, A. L. and Foss, J. F. (eds): 2008, *Handbook of experimental fluid mechanics*, Springer, Berlin.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y.: 2005, A tale of two time scales: determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association* **100**(472), 1394–1411.