WIND
ENERGY
SCIENCE

eawe
european academy of wind energy

Open Access

Multipoint reconstruction of wind speeds

Christian Behnken, Matthias Wächter, and Joachim Peinke

Institute of Physics/ForWind, University of Oldenburg, Oldenburg, Germany

Correspondence: Joachim Peinke (peinke@uni-oldenburg.de)

Abstract. The most intermittent behaviour of atmospheric turbulence is found for very short timescales. Based on a concatenation of conditional probability density functions (cpdf's) of nested wind speed increments, inspired by a Markov process in scale, we derive a short-time predictor for wind speed fluctuations around a non-stationary mean value and with a corresponding non-stationary variance. As a new quality this short-time predictor enables a multipoint reconstruction of wind data. The used cpdf's are (1) directly estimated from historical data from the offshore research platform FINO1 and (2) obtained from numerical solutions of a family of Fokker–Planck equations in the scale domain. The explicit forms of the Fokker–Planck equations are estimated from the given wind data. A good agreement between the statistics of the generated and measured synthetic wind speed fluctuations is found even on timescales below 1 s. This shows that our approach captures the short-time dynamics of real wind speed fluctuations very well. Our method is extended by taking the non-stationarity of the mean wind speed and its non-stationary variance into account.

1 Introduction

The transition of our energy system, formerly strongly relying on gas and coal, to a decarbonised one, mainly based on wind, solar and hydropower, is still ongoing work, but great progress has been made. From 2005 to 2017 the share in installed capacity of wind (solar) has been increased from 6 % (0.3 %) to 18 % (11.5 %) in the European Union (WindEurope, 2018). The downside of this increasing share of fluctuating renewable energy sources is their integration into the power grid. By analysing measurements of fed-in wind and solar power, it could be shown that their fluctuations strongly deviate from Gaussian behaviour on timescales ranging from hours to seconds (Anvari et al., 2016) and for wind power even for scales below 1 s (Haehne et al., 2018). This survival of the atmospheric intermittency in the power grid poses the grid operators with the great challenge to ensure stable power supply, even under highly volatile conditions. Within this context the term intermittency is used in the spirit of Kolmogorov (1962) to describe the characteristic heavy-tailed shape of pdf's often found at small scales in time series of turbulent systems (Frisch, 2004).

To aid the design of our future energy systems, for example to size the needed energy storage or the power generation capacity of conventional power plants, much work has been done in the field of long-term wind speed and power modelling, utilising Markov chain models. Whereas simple first-order Markov chain models cannot grasp the characteristics of long-term correlations of wind speeds (Brokish and Kirtley, 2009), higher-order Markov chain models perform better, but will require more input data for estimating the transition matrices or some simplifications (Pesch et al., 2015; Brokish and Kirtley, 2009; Papaefthymiou and Klockl, 2008; Weber et al., 2018).

Despite their dramatic effect of long-range correlation and fluctuations of wind speeds on the power generation (and thus the grid stability), wind speed fluctuations are known to be most intermittent on short timescales (Boettcher et al., 2003). With short timescales we refer to timescales in the range of seconds to minutes. As can be seen in Boettcher et al. (2003), the effect of intermittency is most prominent at timescales < 1 s, but as the timescales increase, the pdf's broaden. Models considering time steps ranging from minutes to seconds or even below are of course not suited for energy system analysis on national levels, but they are useful

tools for wind turbine operators. For a time step of 10 min Carpinone et al. (2010) presented a higher-order Markov chain model for wind power, and Milan et al. (2013) showed a stochastic power model based on a conditional Langevin equation to work even in the range of seconds.

The knowledge of full three-dimensional wind fields for all three velocity components and the pressure in all details would be desirable. The lack of basic understanding, the impracticability of handling such huge data sets and the complexity of the wind energy conversion process often lead to the demand for simplified models for wind speed. Common approaches for the design of wind turbines are the so-called Mann uniform shear and the Kaimal spectral and exponential coherence model (IEC, 2005). Both models take spectral and coherence aspects of turbulent velocity fluctuations into account, thus handling the fluctuations as Gaussian distributed and stationary. Higher-order statistical effects like the prominent intermittency effect of turbulence and non-stationarities are not taken into account; see for example Mücke et al. (2011). Another approach is to use one-dimensional effective wind speed time series, representing for example the wind field together with the rotor aerodynamics as it impacts the drivetrain or can be used to model the above-discussed energy conversion process (Wächter et al., 2011).

Within this work we propose a novel stochastic generator of one-dimensional wind speed fluctuations with a sampling interval of 0.1 s. One main novelty is that we show how to grasp by this model multipoint statistics of wind structures in time. While commonly applied methods, like spectral analysis and two-point correlations, limit themselves to two-point statistics, here we extend the methodology to more than two points in time. We obtain generalised correlations between multiple points in time, in terms of probability density functions (pdf's), for the occurrence of a whole sequence of wind speeds. Those pdf's we denote multipoint pdf's, and they constitute the basic concept of our approach. Such a stochastic multipoint approach should in principle be able to grasp wind structures like gusts as well as clustering of wind fluctuations. The method was initially developed by Nawroth and Peinke (2006) in the context of homogeneous isotropic turbulence and later on applied to the modelling of log-return rates of current exchange rates (Nawroth et al., 2010) and velocity increments of idealised homogeneous isotropic turbulence (Stresing and Peinke, 2010). The successful application to ocean gravity waves (Hadjihosseini et al., 2016) showed that structures of monster waves can be grasped by this approach correctly (Hadjihoseini et al., 2018). For a recent review see (Peinke et al., 2019). Finally we want to point out that we also show how to handle the aspect of non-stationary wind conditions.

We will continue as follows. In the first part we discuss the method for a subset of wind data characterised by its mean wind speed and its standard deviation. For such data it is shown how, arising from a Langevin process in scale, a predictor for the upcoming wind speed fluctuation around a

mean value can be derived by a nesting of conditional probability density functions. Afterwards we check for Markovian properties of the wind speed fluctuations in scale and set up a Fokker–Planck equation, corresponding to the Langevin process in scale, and we show how it contributes to the improvement of our stochastic prediction method. Finally, in the second part, the non-stationary mean wind speed and its non-stationary variance are incorporated into our approach to achieve more realistic wind speed time series.

2   Method

In this section we present the stochastic framework used for our multipoint reconstruction scheme. As a simplification we start this discussion for blocks of wind data $U(t)$, with the time t, which share the same mean wind speed $\overline{U}$ and the same standard deviation σ_U, as suggested in Morales et al. (2012). Fixing $\overline{U}$ and σ_U one would generate quasi-stationary subsets of data. We follow this idea, but instead of creating quasi-stationary subsets, we aim at creating quasi-stationary time series. With $U(t)$ we refer to the resulting wind speed from the horizontal components. The quasi-stationary wind speed $u(t)$ is then obtained from $U(t)$ by normalising it with the mean $\overline{U}$ and standard deviation σ_U within blocks of 1 min length. We use measured data from the offshore research platform FINO1: the data were recorded at a sampling frequency of 1 Hz between calendar weeks 1 and 10 in 2007 with an ultrasonic anemometer, mounted at 80 m height, resulting in approximately 6×10^6 samples. The use of our method for non-stationary time series is outlined in Sect. 3.

2.1   Multipoint statistics

Since we assume wind speeds to emerge from a turbulent cascade, increments will play a key role in our method. Having a time series of wind speeds $u(t)$, the corresponding increment time series $\Delta u(\tau)$, depending on a certain scale τ, is given by

$$\Delta u(\tau) = u(t) - u(t - \tau). \tag{1}$$

This is the definition of so-called right-sided increments. Note that the calculation of $\Delta u(\tau)$ after this definition depends on the current wind speed $u(t)$ and a past value $u(t-\tau)$, whereas the use of left-sided increments $\Delta u(\tau) = u(t+\tau) - u(t)$ would premise the knowledge of future values $u(t+\tau)$, creating a contradiction as we aim at producing synthetic wind speed time series. To ease readability we use the shorthand notations $u_i := u(t-\tau_i)$ and $\Delta u_i := u(t) - u(t-\tau_i)$ with $\tau_i = i \cdot \tau$ $(i = 1, 2, 3, \ldots)$ in the following.

As a further remark we note that although we consider in this work time series of wind speed, we often talk of multipoint statistics. $\Delta u(\tau)$ is considered to be a statistical two-point quantity, which more correctly could be denoted as a

two-time quantity. Based on the commonly used hypotheses on frozen turbulence by Taylor, for short-time fluctuations time and space statistics are related linearly by the mean wind speed (see also Peinke et al., 2019).

Our idea is to predict a wind speed $u^*(t^*)$ only by knowledge of its N past values $\{u_1(t^* - \tau_1), \ldots, u_N(t^* - \tau_N)\}$. The probability of an event u^* to happen at time t^* can then be expressed by the conditional probability density function (cpdf)

$$p(u^*, t^* | u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N) =$$
$$\frac{p(u^*, t^*; u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N)}{p(u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N)}, \quad (2)$$

using the definition of conditional probabilities. Note this cpdf is the key quantity to estimate a new wind speed value u^*, and it can be used iteratively to generate new time series, as we show below.

Now we link Eq. (2) to the idea of an underlying turbulent cascade, which can be described by means of a Markov process. Thus we assume that there exists a scale separation $\Delta \tau_{ME} = \tau_j - \tau_i$ $(j > i)$, after which the Markov condition is fulfilled for all greater scales. This scale separation $\Delta \tau_{ME}$ is often called the Markov–Einstein length (Einstein, 1905). It enables us to express the multiple cpdf $p(u^*, t^* | u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N)$ by a multiplication of a simpler double cpdf and marginal pdf's:

$$p(u^*, t^* | u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N) =$$
$$\frac{p(u^*)}{p(u_1)} \cdot \frac{\Pi_{i=1}^{N-1} p(\Delta u_i | \Delta u_{i+1}; u^*)}{\Pi_{i=2}^{N-1} p(\tilde{\Delta u_i} | \tilde{\Delta u_{i+1}}; u_1)} \cdot \frac{p(\Delta u_N | u^*)}{p(\tilde{\Delta u_N} | u_1)}, \quad (3)$$

with $\tilde{\Delta u_i} := u(t^* - \tau_1) - u(t^* - \tau_i)$ with the timescale $\tau_i - \tau_1$. For details we refer the reader to Appendix A.

It is known that the evolution of cpdf's of a Markov process can be described by the famous Kramers–Moyal expansion (Risken, 1996), which notes for our stochastic process in scale Δu_i

$$-\tau_i \frac{\partial}{\partial \tau_i} p(\Delta u_i | \Delta u_j; u^*) =$$
$$\sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial \Delta u_i} \right)^n \left[D^{(n)}(\Delta u_i, \tau_i, u^*) p(\Delta u_i | \Delta u_j; u^*) \right], \quad (4)$$

requiring $\tau_j - \tau_i \geq \Delta \tau_{ME}$ and with the Kramers–Moyal coefficients $D^{(n)}$ being defined as

$$D^{(n)}(\Delta u_i, \tau_i, u^*) =$$
$$\lim_{\Delta \tau \to 0} \frac{\tau_i}{n! \Delta \tau} \langle \left[\Delta u_i'(\tau_i - \Delta \tau, u^*) - \Delta u_i(\tau_i, u^*) \right]^n \rangle. \quad (5)$$

In contrast to the Kramers–Moyal expansion in time domain, a minus sign on the left-hand side (lhs) of Eq. (4) has to be added for scale processes, since during evolution of the process the scale τ is decreasing. According to the Pawula theorem, the Kramers–Moyal coefficients vanish for $n \geq 3$, if

$D^{(4)} = 0$ (Risken, 1996); thus the Kramers–Moyal expansion reduces to the Fokker–Planck equation (FPE):

$$-\tau_i \frac{\partial}{\partial \tau_i} p(\Delta u_i | \Delta u_j; u^*) =$$
$$-\frac{\partial}{\partial \Delta u_i} \left[D^{(1)}(\Delta u_i, \tau_i, u^*) p(\Delta u_i | \Delta u_j; u^*) \right] +$$
$$-\frac{\partial^2}{\partial \Delta u_i^2} \left[D^{(2)}(\Delta u_i, \tau_i, u^*) p(\Delta u_i | \Delta u_j; u^*) \right], \quad (6)$$

with the drift function $D^{(1)}$ accounting for the deterministic evolution of the stochastic process, whereas the so-called diffusion function $D^{(2)}$ scales the amplitude of the noise term of the corresponding Langevin equation

$$-\frac{\partial}{\partial \tau} \Delta u(\tau, u^*) =$$
$$\frac{1}{\tau} D^{(1)}(\Delta u, \tau, u^*) + \sqrt{\frac{1}{\tau} D^{(2)}(\Delta u, \tau, u^*)} \cdot \Gamma(\tau) \quad (7)$$

with the Gaussian noise $\Gamma(\tau)$, fulfilling $\langle \Gamma(\tau) \rangle = 0$ and also $\langle \Gamma(\tau) \Gamma(\tau') \rangle = 2 \delta_{\tau \tau'}$. This equation directly describes the evolution of a single trajectory along the scale τ.

As can be seen, the FPE can be used to determine the factorised pdf's from Eq. (3) if the process is Markovian and higher-order Kramers–Moyal coefficients are zero. This equivalence between a cpdf with N conditions and a nested chain of several cpdf's with only two conditions, stemming from the three-point closure of a cascade process, is extremely helpful if one aims at estimating the pdf's from measurements, since the high-dimensional pdf's in Eq. (2) would require a tremendous number of realisations in order to be estimated well. For a more detailed discussion of this multipoint approach we refer to Peinke et al. (2019).

2.2 Preliminary analysis of wind speed data

Next we check if wind speed data are suitable for the reconstruction method just described. According to the right-hand side (rhs) of Eq. (3) the estimation of the double cpdf's $p(\Delta u_i | \Delta u_{i+1}, u^*)$ is necessary. However, to reduce computational costs or the number of data points, it would be much more convenient to use the single cpdf's $p(\Delta u_i | \Delta u_{i+1})$ by excluding the condition on the wind speed u^* to be predicted (Nawroth et al., 2010; Peinke et al., 2019). Thus the equality

$$p(\Delta u_i | \Delta u_{i+1}; u^*) = p(\Delta u_i | \Delta u_{i+1}) \quad (8)$$

must hold. As can be seen from Fig. 1, Eq. (8) holds for $u^* \approx 0$ but shows a significant shift for $u^* \approx 2.5$. From this we reason that the equality in Eq. (8) cannot generally be assumed, so we have to stick to the double cpdf's $p(\Delta u_i | \Delta u_{i+1}, u^*)$. Similar results have been reported for idealised turbulence (Stresing and Peinke, 2010) and sea waves (Hadjihosseini et al., 2016). On derivation of our final formula for the reconstruction of time series (see Eq. 3) an essential step was to

premise the underlying scale process to be Markovian. Thus it has to be shown that for $\Delta\tau = \tau_{i+1} - \tau_i \geq \Delta\tau_{\mathrm{ME}}$

$$p(\Delta u_i | \Delta u_{i+1}; \ldots; \Delta u_N; u^*) = p(\Delta u_i | \Delta u_{i+1}; u^*) \qquad (9)$$

is a valid assumption. As the verification of this expression is not feasible in its generality, we limit ourselves by reducing the number of dimensions involved and just check the equality of

$$p(\Delta u_i | \Delta u_{i+1}; \Delta u_{i+2}; u^*) = p(\Delta u_i | \Delta u_{i+1}; u^*). \qquad (10)$$

To check this we utilise the Chapman–Kolmogorov equation (CKE) (Friedrich et al., 2011). The cpdf $p(\Delta u_i | \Delta u_{i+2}; u^*)$ is estimated directly from observational data and afterwards compared with the cpdf $\widetilde{p}(\Delta u_i | \Delta u_{i+2}; u^*)$ obtained numerically by use of the CKE

$$\tilde{p}(\Delta u_i | \Delta u_{i+2}; u^*) =$$
$$\int p(\Delta u_i | \Delta u_{i+1}; u^*) \, p(\Delta u_{i+1} | \Delta u_{i+2}; u^*), \qquad (11)$$

whereas the two cpdf's within the integral on the rhs are also directly estimated from data. Figure 2 shows that Eq. (11) only holds for $\Delta\tau \geq \Delta\tau_{\mathrm{ME}}$. We find a Markov–Einstein length of $\Delta\tau_{\mathrm{ME}} \leq 0.1\,\mathrm{s}$, which we are going to use henceforth.

2.3 Parameterisation of the Fokker–Planck equation

As was mentioned in Sect. 2.1 the FPE may be used to generate solutions for the needed cpdf's in Eq. (1). Aiming for this, one needs a parameterisation of the FPE reflecting the scale process of the real-world wind speed data. No general, physical formulation for a FPE describing the scale process of wind speeds is known; thus we use the possibility to estimate a parameterisation directly from the given data (compare Eq. 5 and Reinke et al., 2018; Peinke et al., 2019).

This way we obtain estimations of the drift and diffusion functions $D^{(1)}(\Delta u, \tau_i, u^*)$, $D^{(2)}(\Delta u, \tau_i, u^*)$ along the scale τ_i and for every wind speed value u^*. From these estimations one then usually finds parameterisations of the FPE by fitting appropriate polynomial surfaces to the estimated functions, which then may be used to obtain numerical solutions of the FPE.

To match the functional shape of the estimated $D^{(1)}(\Delta u, \tau_i, u^*)$ and $D^{(2)}(\Delta u, \tau_i, u^*)$ (see Fig. 3), we require the polynomials to be of the order of 3 and 2 respectively. We find a significant shift, denoted with $\gamma_{D^{(1)}}(\tau_i, u^*)$, depending on u^* and τ_i for the drift along all scales τ_i which has to be taken into account for a parameterisation suitable to the given data.

We set for $D^{(1)}(\Delta u, \tau_i, u^*)$

$$D^{(1)}(\Delta u, \tau_i, u^*) =$$
$$d_{10}(\tau_i, u^*) + d_{11}(\tau_i, u^*)\big[u - \gamma_{D^{(1)}}(\tau_i, u^*)\big] +$$
$$d_{13}(\tau_i, u^*)\big[u - \gamma_{D^{(1)}}(\tau_i, u^*)\big]^3 \qquad (12)$$

and for $D^{(2)}(\Delta u, \tau_i, u^*)$

$$D^{(2)}(\Delta u, \tau_i, u^*) =$$
$$d_{20}(\tau_i, u^*) + d_{21}(\tau_i, u^*)u + d_{22}(\tau_i, u^*)u^2. \qquad (13)$$

Similar findings were obtained for other systems (Hadjihosseini et al., 2016; Stresing and Peinke, 2010), where the dependence on u^* was limited only to the drift function $D^{(1)}(\Delta u, \tau_i, u^*)$. For wind speed data the contribution of u^* is clearly more complex.

Furthermore we check the validity of the Pawula theorem, requiring $D^{(4)} = 0$. As one can see in Fig. 4, the fourth Kramers–Moyal coefficient is slightly larger than zero, but negligible compared to the magnitude of the diffusion function $D^{(2)}$.

Next, we check if the parameterisation of the FPE, given by Eqs. (12) and (13), performs well in describing the underlying scale process, before one uses it for the reconstruction scheme presented in Sect. 2.1. This can easily be done by comparing cpdf's estimated directly from the data and the ones obtained from numerical solutions of the FPE. The latter one is not carried out by common finite-difference schemes but by an iterative approach (for first works see Renner et al., 2001; Wächter et al., 2003). For a (very) small step size in scale $\Delta\tau$ the functions $D^{(1)}(\Delta u, \tau_i, u^*)$ and $D^{(2)}(\Delta u, \tau_i, u^*)$ can be assumed to be constant in τ, leading to an exact solution (see Risken, 1996) for the cpdf

$$p(\Delta u_j, \tau_i - \Delta\tau | \Delta u_i, \tau_i, u^*) = \frac{1}{2\sqrt{\pi D^{(2)}(\tau_i, u^*)\Delta\tau}} \cdot$$
$$\exp\left[-\frac{\big(\Delta u_j - \Delta u_i - D^{(1)}(\tau_i, u^*)\Delta\tau\big)^2}{4 D^{(2)}(\tau_i, u^*)\Delta\tau} \right], \qquad (14)$$

describing the transition between wind speed increments of a larger scale τ_i and a smaller scale $\tau_i - \Delta\tau$. By iteratively combining this so-called short-time propagator (STP) with the CKE (see Eq. 11), one is able to obtain cpdf's $p(\Delta u_j, \tau_j | \Delta u_i, \tau_i, u^*)$ for arbitrary large-scale differences $\tau_i - \tau_j \gg \Delta\tau$. In theory this procedure would lead to exact solutions of the FPE, but since one is limited to finite step sizes $\Delta\tau$, this method of course only provides numerical approximations of the true cpdf's.

Comparing cpdf's estimated directly from the data and from the numerical solution, we see (Fig. 5) that our proposed estimation in terms of $D^{(1)}(\Delta u, \tau_i, u^*)$ and $D^{(2)}(\Delta u, \tau_i, u^*)$ is well suited to describe the underlying scale process. Thus we use this parameterisation for the reconstruction scheme.
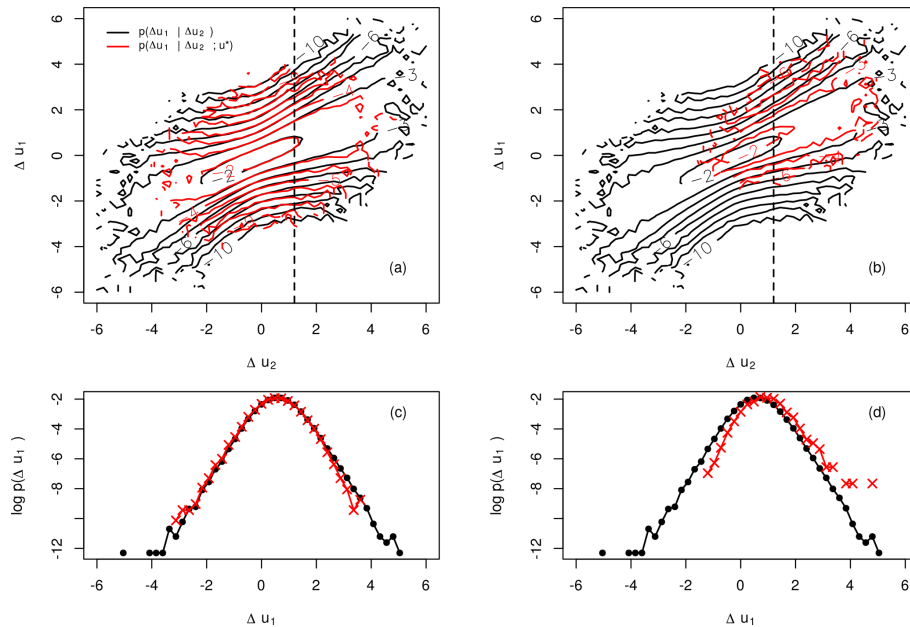
Figure 1. Comparison of single cpdf's $p(\Delta u_1|\Delta u_2)$ (black) and double cpdf's $p(\Delta u_1|\Delta u_2; u^*)$ (red) for $u^* \approx 0$ (**a, c**), $u^* \approx 2.5$ (**b, d**), $\tau_1 = 1\,\text{s}$ and $\tau_2 = 2\tau_1$. Dashed lines indicate cuts through the cpdf's at $\Delta u_2 \approx 1.2$. The levels of the contour plots are given in natural logarithmic scale.
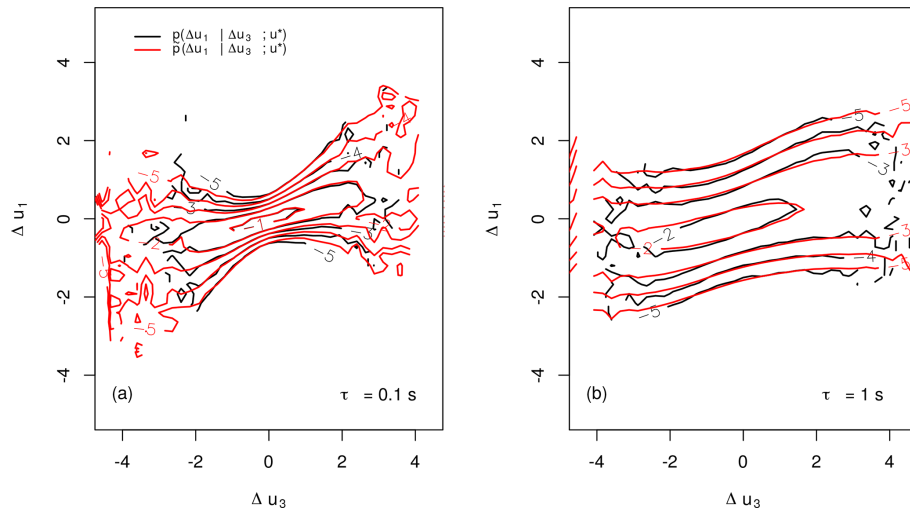

Figure 2. Comparison of the double cpdf's $p(\Delta u_1|\Delta u_3; u^*)$ (estimated from data) and $\widetilde{p}(\Delta u_1|\Delta u_3; u^*)$ (obtained from CKE) for $\tau_1 = 0.1\,\text{s}$ (**a**) or $1\,\text{s}$ (**b**) and $\tau_2 = 2\tau_1$, $\tau_3 = 3\tau_1$, $\Delta \tau = \tau_2 - \tau_1 = \tau_3 - \tau_2$ and $u^* = 0$.

2.4 Results of the multiscale reconstruction

Alternatively to the presented approach to obtain the cpdf's from numerical solutions of the FPE, it is of course possible and much less cumbersome to estimate them directly from observational data. (Note that due to the use of the FPE, the obtained pdf's are less noisy and extend to large values as seen in Fig. 5.) In this section we will present the results for the multipoint reconstruction achieved for u^*, yielded from both approaches.

To start the reconstruction scheme, we provided a short piece of the original time series of N wind speeds to the algorithm to compute the pdf $p(u^*|u_1, \ldots, u_N)$, from which the first point of our simulation can be drawn. The reconstruction scheme is then shifted by one time step τ and applied again. By iteratively applying our method a new artificial time series of arbitrary length can be generated. After N iterations of the reconstruction scheme no data from measurements are required any more to generate new values of u^*. From visual comparison in Fig. 6 one finds a realistic looking simulated
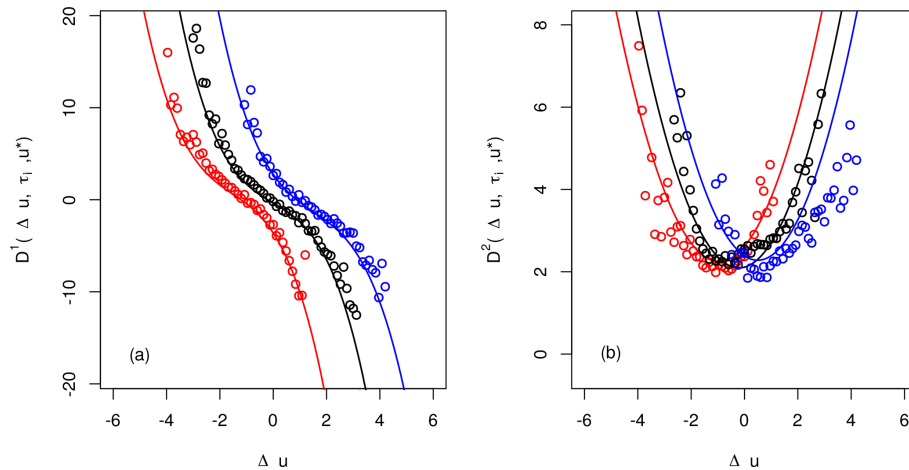
Figure 3. Exemplary estimations of the drift and diffusion functions $D^{(1)}(\Delta u, \tau_i, u^*)$ and $D^{(2)}(\Delta u, \tau_i, u^*)$ for $\tau = 65\,\mathrm{s}$ and $u^* \approx -1.58$ (red), ≈ 0 (black) and ≈ 1.58 (blue). Fits for $D^{(1)}$ and $D^{(2)}$ were achieved by applying Eqs. (12) and (13).
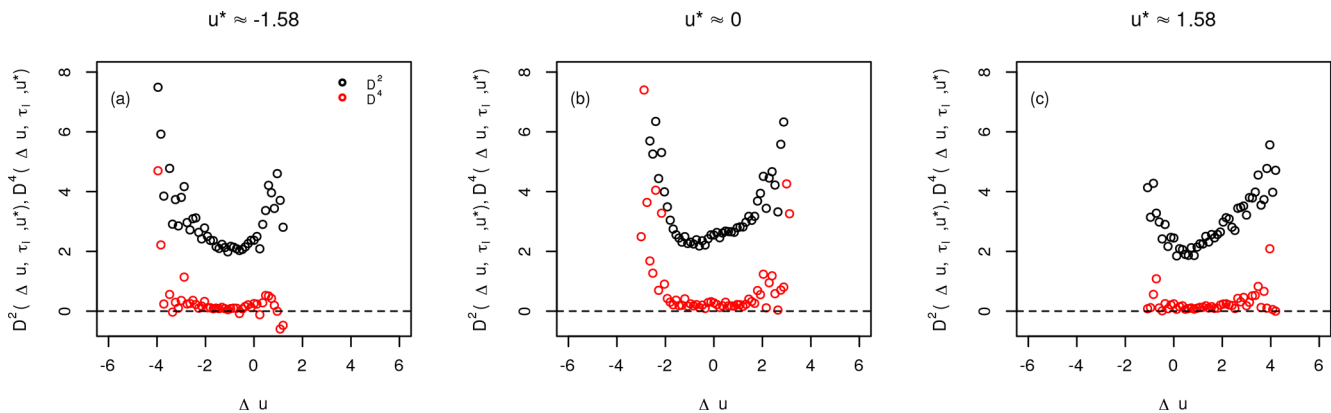

Figure 4. Exemplary estimations of the second and fourth Kramers–Moyal coefficients $D^{(2)}(\Delta u, \tau_i, u^*)$ and $D^{(4)}(\Delta u, \tau_i, u^*)$ for $\tau = 65\,\mathrm{s}$.

time series of u^*, retaining the characteristics of dynamics on the smaller scales as well as the ones of the larger scales.

To confirm this visual impression in a quantitative way, we compare the increment pdf's $p(\Delta u_i, \tau_i)$ obtained from the reconstructed and the measured time series. The synthetic time series were generated by using the cpdf's estimated directly both from data and from numerical solutions of the FPE.

As shown in Fig. 7 the increment pdf's yielded from both stochastic simulations nicely coincide with the pdf's from observations. This attests that the presented reconstruction scheme is able to capture the complex dynamics of wind speeds, characterised by a gradual shift of increment pdf's of a Gaussian-like shape (larger scales) to increment pdf's of a heavy-tailed shape (smaller scales).

A striking difference between the increment pdf's from the stochastic simulation can be noted. Whereas the tails of the original pdf's are systematically underestimated by the reconstruction using the directly estimated pdf's, the reconstruction from the numerically obtained cpdf's is able to keep track of the tails of the original pdf. This stems from the fact,

mentioned above, that by estimating a pdf from observational data one in general underestimates the outer tails, since there are only a few measured points available. Considering the estimation of cpdf's, the estimation error of course worsens, which is even more severe when additional conditioning is applied, like in our case in terms of the additional condition on u^*.

The tails of cpdf's $p(\Delta u_i, \tau_i | \Delta u_j, \tau_j; u^*)$ computed from the family of FPEs can be extrapolated to areas where no measurement points are available. This effect is a direct consequence of the CKE (Eq. 11), as the tails are the product of quite well-estimated less probable but not rare events.

Certainly this approach indirectly suffers from the limited number of observations as well, as the estimation of the functions $D^{(1)}(u, r, u^*)$ and $D^{(2)}(u, r, u^*)$ is based on observational data, too (see Eq. 5). Furthermore the parameterisation of these functions is always only a approximation of the real drift and diffusion functions, introducing deviations from the real-world system.
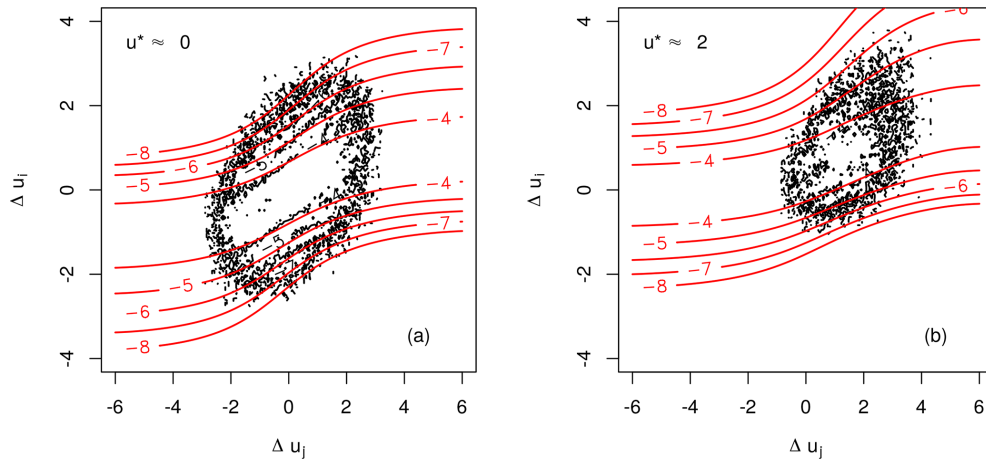
Figure 5. Isoline plots of a cpdf $p(\Delta u_i, \tau_i | \Delta u_j, \tau_j; u^*)$ estimated from the data (black) and from the numerical solution of the FPE (red) for $\tau_i = 1.6\,\mathrm{s}$ and $\tau_j = 3.2\,\mathrm{s}$ and for $u^* \approx 0$ and $u^* \approx 2$.
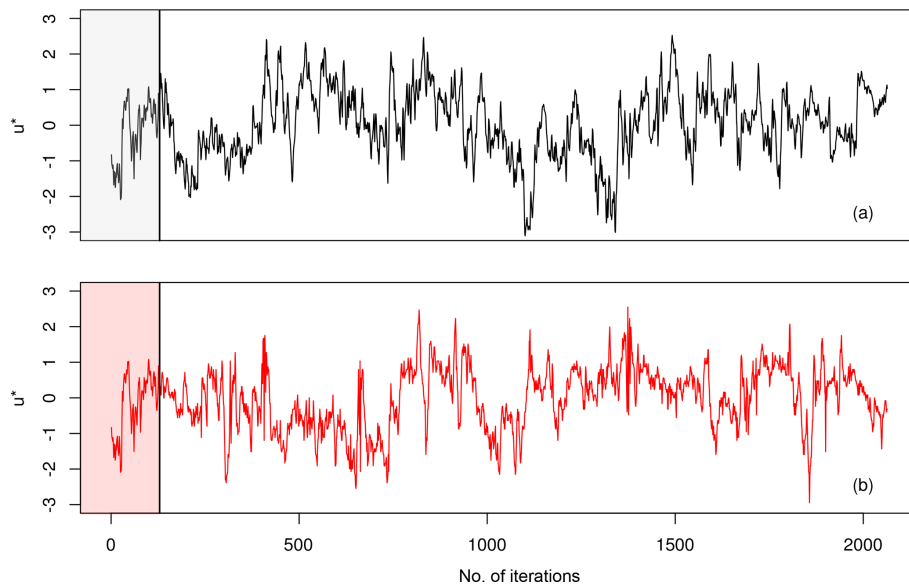

Figure 6. Comparison of original wind speed time series u^* (black, **a**) with the reconstructed one using Eq. (8) (red, **b**). The vertical line marks the transition from the $N = 128$ starting values to the simulated wind speeds.

But we see from Fig. 7 that the majority of increments are well grasped; only the occurrence of a few rare events are underpredicted. A more detailed investigation of the pdf's shows the pdf's obtained by the FPE deviate from the original shape for the largest scale $\tau = N \cdot \Delta \tau_{\mathrm{ME}}$ but performs better in the outer wings of the pdf.

From Fig. 8 a better understanding of the reconstruction method can be gained. The pdf used for the simulation $p(u^* | u_1, \dots, u_N)$ is not stationary, even though it is computed from completely stationary cpdf's (see Eq. 8) and changes sensitively with respect to the N past values. While the snapshot pdf (shown as black circles) at the time marked by the black vertical line has a rather clear shape, it undergoes a change, becoming broader (red line and red circles). Due to

the spreading of the wings of the pdf, values of u^* of larger magnitude become more likely to be drawn. This may lead to a distinct shift of the wind speed values to $u^* \leq 0$ as seen in this example. After this transition the broadness decreases (blue line and circles) again and a tendency to relax back to $u^* = 0$ can be seen. Furthermore the increased broadness of the red pdf (red line) can be seen as an early warning signal that the wind speed is prone to fluctuate in a stronger way.

3 Extension to non-stationary wind speeds

In the preceding part for the reconstruction scheme, block-wise normalised wind speeds with a window length of 1 min
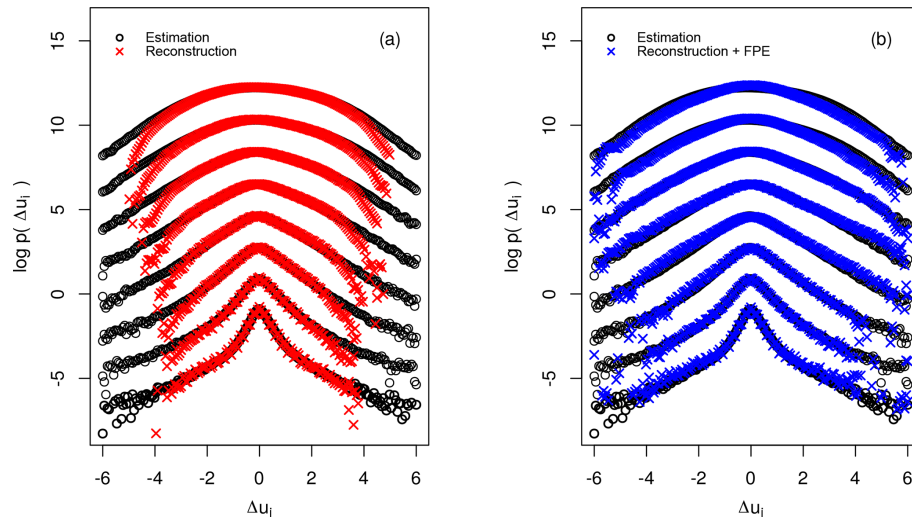
Figure 7. Comparison of the marginal increment pdf's computed from the empirical data (black), the simulated data using the directly estimated pdf's (red) and the simulated data using the cpdf's obtained from numerical solution of the FPE (blue). The scales range from $\Delta \tau_{ME}$ to $N \cdot \tau_{ME}$; they are explicitly: $2^i \cdot \Delta \tau_{ME}$ with $(i = 0, 1, \ldots, 7)$ and $\Delta \tau_{ME} = 0.1$ s. For better visualisation the pdf's were shifted along the vertical axis.
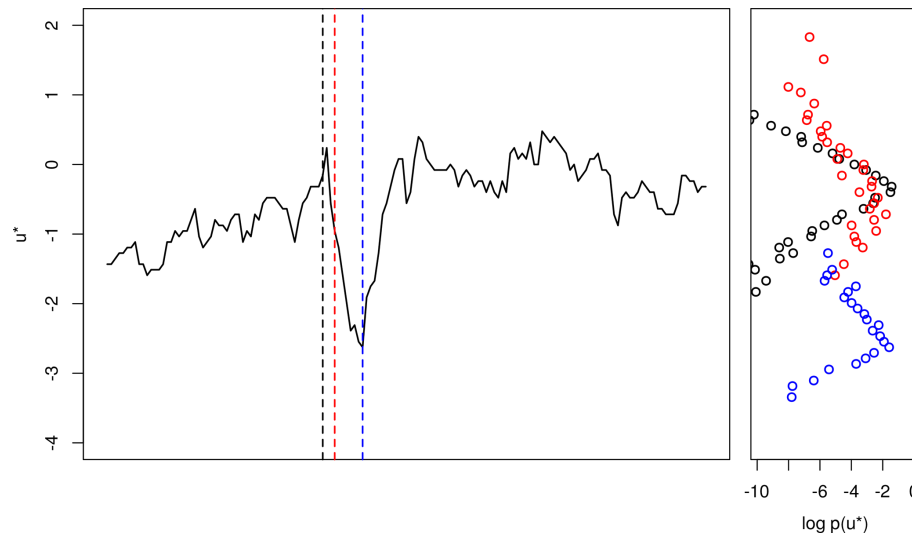
Figure 8. Evolution of $p(u^*|u_1, \ldots, u_N)$ during reconstruction. Horizontal lines indicate snapshots of the pdf used for drawing the next sample of u^*. The colours of the horizontal lines correspond to the snapshot pdf's.

were used. These blocks were defined by common mean wind speed $\overline{U}$ and standard deviation σ_U. For the normalised wind speed u, we showed how to generate new time series; see Fig. 9.

There are different methods to generate more general non-stationary wind data. Knowing the slow variation in $\overline{U}(t)$ and $\sigma_U(t)$, the drift and diffusion coefficients $D^{(i)}$ are taken as a slowly changing function of $D^{(i)}(\Delta u, \tau, u^*, \overline{U}, \sigma_U)$. If due to the normalisation of U to u the coefficients $D^{(i)}$ are in a good approximation independent of $\overline{U}$ and σ_U, the slow variation in the real wind conditions can be incorporated over the back-transformation of the newly estimated values $U^* =$ $(\sigma_U \cdot u^*) + \overline{U}$. The slow dynamics of $\overline{U}(t)$ and $\sigma_U(t)$ may be given from measured data, meteorological simulations or other modelling.

A third possibility is a self-adaptive procedure which we show here. Instead of using given values of $\overline{U}(t)$ and $\sigma_U(t)$, the intrinsic fluctuation of these quantities are used. Given an initial pair $(\overline{U}(t), \sigma_U(t))$ estimated over 1 min from measured data, a time series of non-stationary wind speeds $U^*(t)$ is obtained by applying the above-mentioned back-transformation to the generated wind speeds u^* from our algorithm. For the upcoming simulation window of 1 min length, a new pair $(\overline{U}(t'), \sigma_U(t'))$ is estimated from the just generated block of
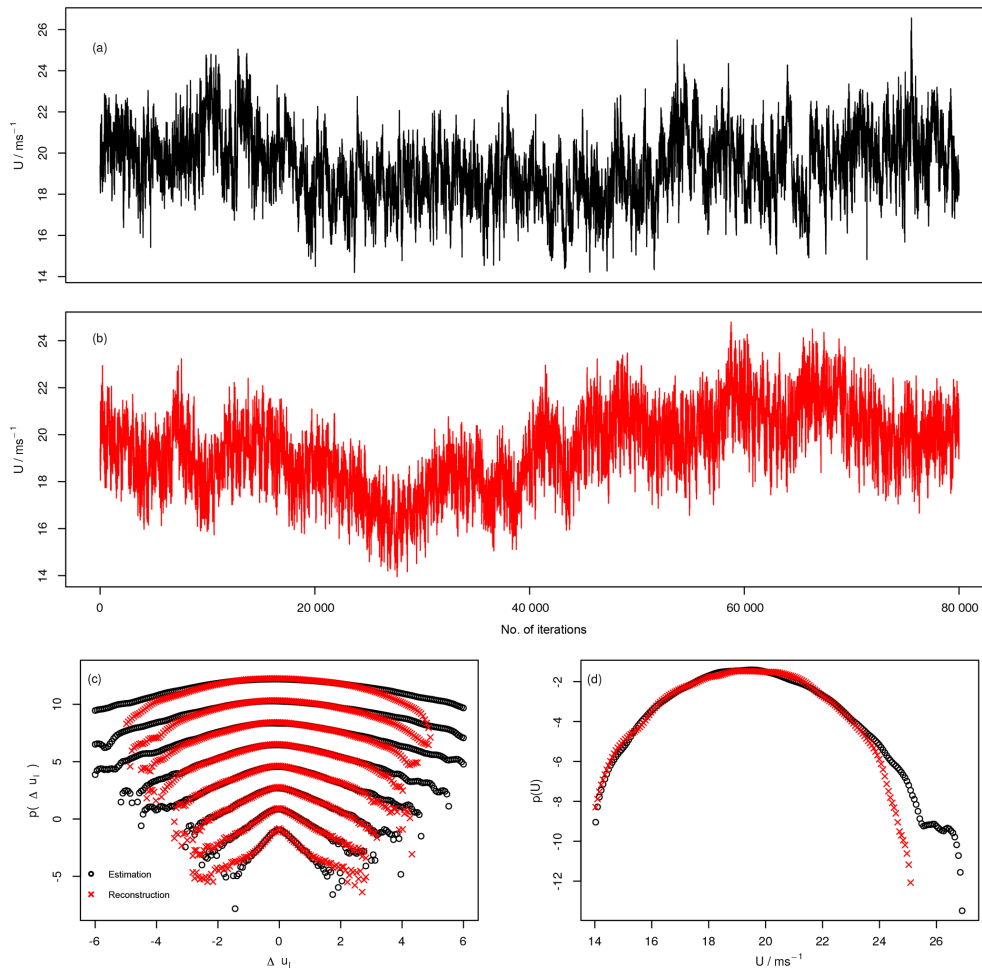
Figure 9. (a, b) Comparison between measured non-stationary speeds U and reconstructed non-stationary wind speeds. **(c, d)** Increment and one point pdf's of original and reconstructed U.

wind speed data $U^*(t)$. This procedure is carried out until a time series of non-stationary wind speeds of desired length is obtained. In Fig. 9 such a time series is shown together with the statistical analysis of the increment pdf's and the marginal pdf of the non-stationary wind speed U. We observe that the pdf's of the reconstructed time series match the shape of the empirical pdf's for both the increments and the wind speeds. At this point we would like to emphasise that we do not aim to create copies of historical wind speeds, but rather to be able to generate stochastically equivalent time series.

4   Conclusions

We presented a stochastic approach based on multipoint statistics to generate surrogate short-time wind speed fluctuations with stochastic processes. Note these stochastic processes can be estimated self-consistently from given data. By using the normalised wind speeds u^* and wind speed increments $\Delta u(\tau_i)$, $\Delta u(\tau_j)$ from two separate scales τ_i and τ_j, a

three-point closure to the complex systems of wind speeds was achieved.

It was shown that our method works well in describing the dynamics of block-wise normalised wind speeds u^* along scales τ_i in terms of a stochastic scale process, governed by a family of Fokker–Planck equations. This separation of the fluctuations from the mean values is similar to the Reynolds-averaged Navier–Stokes (RANS) approach widely used in fluid dynamics (Frisch, 2004), with the difference that we have a description of the underlying stochastic process of the fluctuation and "only" lack the mean values. With the modified reconstruction (see Sect. 3) we are able to generate mean values on the basis of past values of the reconstructed time series, yielding realistic non-stationary wind speed time series U^*. As the typical response times of wind turbines and their control systems have duration of seconds to minutes, our reconstructed wind data are suitable for investigation of many dynamical effects of the wind energy conversion process. Note for these times the wind energy system may be driven into non-stationary response dynamics. Thus for these

timescales it is important to have to maximal information on all statistical details of the driving wind source. The stochastic model presented here is based on multipoint statistics and is able to capture small-scale intermittency effects, extreme events and clustering of fluctuations, up to now not addressed in wind energy research.

For timescales larger than the response times of wind turbines, the turbines operate with fully adapted control systems in a stationary state. To estimate effects, like loads, of such stationary states, the temporal order of the states becomes unimportant. It is sufficient to know how often which wind situation emerges. Thus the knowledge of the valid Weibull distribution $p(\overline{U})$ should be sufficient. Note our result here indicates that it would be better to extend the Weibull distribution to the joint probability $p(\overline{U}, \sigma_U)$.

Finally we emphasise that the presented stochastic multipoint approach to small-scale wind speed fluctuations should encompass automatically extreme short-term wind fluctuations, commonly added to wind investigations in terms of standard or multi-year gusts. These methods can be applied easily to other wind quantities like the temporal behaviour of shears, or wind veers, eventually combined in higher-dimensional stochastic processes (Siefert and Peinke, 2006). The results reported in Ali et al. (2019) show that such a stochastic modelling can also be used for wake flows.

Appendix A: Derivation of simpler cpdf's

To link Eq. (2) to the idea of an underlying turbulent cascade, we identify the conditional pdf on the left-hand side (lhs) with the cpdf $p(u^*, t^* | \Delta u_1, t^* - \tau_1; \ldots; \Delta u_N, t^* - \tau_N)$. Thus the numerator of the right-hand side (rhs of Eq. 2) can be rewritten as

$$
\begin{aligned}
p(u^*, t^*; u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N) = \\
p(u^*, u^* - u_1, \tau_1; \ldots; u^* - u_N, \tau_N)
\end{aligned} \tag{A1}
$$

and the nominator as

$$
\begin{aligned}
p(u_1, t^* - \tau_1; \ldots; u_N, t^* - \tau_N) = \\
p(u_1; u_1 - u_2, \tau_2 - \tau_1; \ldots; u_1 - u_N, \tau_N - \tau_1).
\end{aligned} \tag{A2}
$$

The identity of the expressions in Eqs. (A1) and (A2) can mathematically rigorously be shown, as done in Nawroth et al. (2010), but intuitively speaking the sequences on the lhs and rhs must yield the same joint pdf, since the increments on the rhs have a common reference point u^* or u_1. Next we factorise the joint pdf's from Eqs. (A1) and (A2) by iteratively using cpdf's

$$
\begin{aligned}
p(u^*; \Delta u_1, \tau_1; \ldots; \Delta u_N, \tau_N) = \\
p(\Delta u_1, \tau_1 | \Delta u_2, \tau_2; \ldots; \Delta u_N, \tau_N; u^*) \cdot \\
p(\Delta u_2, \tau_2 | \Delta u_3, \tau_3; \ldots; \Delta u_N, \tau_N; u^*) \cdots \\
p(\Delta u_{N-1}, \tau_{N-1} | \Delta u_N, \tau_N; u^*) \cdot \\
p(\Delta u_N, \tau_N | u^*) \cdot p(u^*)
\end{aligned} \tag{A3}
$$

and with $\Delta \tilde{u}_i := u(t^* - \tau_1) - u(t^* - \tau_i)$ with the timescale $\tau_i - \tau_1$:

$$
\begin{aligned}
p(u_1; \Delta \tilde{u}_2, \tau_2 - \tau_1; \ldots; \Delta \tilde{u}_N, \tau_N - \tau_1) = \\
p(\Delta \tilde{u}_2, \tau_2 - \tau_1 | \Delta \tilde{u}_3, \tau_3 - \tau_1; \ldots; \Delta \tilde{u}_N, \tau_N - \tau_1; u_1) \cdot \\
p(\Delta \tilde{u}_3, \tau_3 - \tau_1 | \Delta \tilde{u}_4, \tau_4 - \tau_1; \ldots; \Delta \tilde{u}_N, \tau_N - \tau_1; u_1) \\
\cdots p(\Delta \tilde{u}_{N_1}, \tau_{N-1} - \tau_1 | \Delta \tilde{u}_N, \tau_N - \tau_1; u_1) \cdot \\
p(\Delta \tilde{u}_N, \tau_N - \tau_1 | u_1) \cdot p(u_1).
\end{aligned} \tag{A4}
$$

A further step in reducing the dimensionality of the involved pdf's can be performed upon assuming the scale process to be Markovian; i.e. there exists a timescale separation $\Delta \tau_{ME} = \tau_j - \tau_i$ ($j > i$), where

$$
\begin{aligned}
p(\Delta u_i, \tau_i | \Delta u_j, \tau_i + \Delta \tau_{ME}; \ldots; \Delta u_n, \tau_i + n \cdot \Delta \tau_{ME}; u^*) \\
= p(\Delta u_i, \tau_i | \Delta u_j, \tau_i + \Delta \tau_{ME}; u^*)
\end{aligned} \tag{A5}
$$

holds. The timescale separation $\Delta \tau_{ME}$ is often called the Markov–Einstein length (Einstein, 1905), and for various systems its existence could be shown empirically, ranging from jet streams in laboratory experiments (Renner et al., 2001), (Reinke et al., 2018) to large geophysical systems such as ocean gravity waves (Hadjihosseini et al., 2016).

Appendix B: Parameterisation of $D^{(1)}$ and $D^{(2)}$

Here we present the polynomial coefficients used to parameterise the first and second Kramers–Moyal coefficients $D^{(1)}(\Delta u_i, \tau_i, u^*)$ and $D^{(2)}(\Delta u_i, \tau_i, u^*)$ (see Eqs. 12 and 13).

Coefficients for the drift function $D^{(1)}(\Delta u_i, \tau_i, u^*)$:

$$
\begin{aligned}
d_{10} = c_{0,d_{10}} \cdot \tau_i + c_{1,d_{10}} \cdot u^* \cdot \tau_i^{\tilde{c}_{1,d_{10}}} + c_{2,d_{10}} \cdot \tau_i^{\tilde{c}_{2,d_{10}}} \\
\cdot u^{*2} + c_{3,d_{10}} \cdot u^{*3},
\end{aligned} \tag{B1}
$$

$$
d_{11} = c_{0,d_{11}} \cdot \tau_i + c_{1,d_{11}} \cdot \tau_i^{\tilde{c}_{1,d_{11}}} + c_{2,d_{11}} \cdot \tau_i^{\tilde{c}_{2,d_{11}}} \cdot u^{*2}, \tag{B2}
$$

$$
d_{13} = c_{0,d_{13}} \cdot \tau_i^{\tilde{c}_{0,d_{13}}} + c_{1,d_{13}} \cdot u^*, \tag{B3}
$$

$$
\gamma_{D^{(1)}} = c_{1,\gamma_{D^{(1)}}} \cdot \tau_i^{\tilde{c}_{1,\gamma_{D^{(1)}}}} \cdot u^*, \tag{B4}
$$

with $c_{0,d_{10}} = -0.006$, $c_{1,d_{10}} = -0.888$, $\tilde{c}_{1,d_{10}} = 0.098$, $c_{2,d_{10}} = 0.137$, $\tilde{c}_{2,d_{10}} = 0.019$, $c_{3,d_{10}} = -10.566$, $c_{0,d_{11}} = -1.656$, $c_{1,d_{11}} = -0.018$, $\tilde{c}_{1,d_{11}} = -8.853e - 05$, $c_{2,d_{11}} = -0.268$, $\tilde{c}_{2,d_{11}} = 1.671$, $c_{0,d_{13}} = -0.005$, $\tilde{c}_{0,d_{13}} = 0.012$, $c_{1,d_{13}} = 1.023$, $c_{1,\gamma_{D^{(1)}}} = 0.341$, $\tilde{c}_{1,\gamma_{D^{(1)}}} = 0.247$.

And for the diffusion function $D^{(2)}(\Delta u_i, \tau_i, u^*)$

$$
\begin{aligned}
d_{20} = c_{0,d_{20}} \cdot \tau_i^{\tilde{c}_{0,d_{20}}} + c_{1,d_{20}} \cdot \tau_i^{\tilde{c}_{1,d_{20}}} \cdot u^* + c_{2,d_{20}} \\
\cdot \tau_i^{\tilde{c}_{2,d_{20}}} \cdot u^{*2},
\end{aligned} \tag{B5}
$$

$$
d_{21} = c_{0,d_{21}} \cdot \tau_i^{\tilde{c}_{0,d_{21}}} + c_{1,d_{21}} \cdot \tau_i^{\tilde{c}_{1,d_{21}}} \cdot u^*, \tag{B6}
$$

$$
d_{22} = c_{0,d_{22}} \cdot \tau_i^{\tilde{c}_{0,d_{22}}} + c_{1,d_{22}} \cdot \tau_i^{\tilde{c}_{1,d_{22}}} \cdot u^*, \tag{B7}
$$

with $c_{0,d_{20}} = 0.024$, $\tilde{c}_{0,d_{20}} = -0.0001$, $c_{1,d_{20}} = 0.0002$, $\tilde{c}_{1,d_{20}} = 1.076$, $c_{2,d_{20}} = 1.573$, $\tilde{c}_{2,d_{20}} = 1.622$, $c_{0,d_{21}} = 0.002$, $\tilde{c}_{0,d_{21}} = -0.001$, $c_{1,d_{21}} = 1.104$, $\tilde{c}_{1,d_{21}} = 1.395$, $c_{0,d_{22}} = 0.042$, $\tilde{c}_{0,d_{22}} = 0.002$, $c_{1,d_{22}} = 0.555$, $\tilde{c}_{1,d_{22}} = 0.364$.

References

IEC: IEC 61400-1 Wind turbines Part 1: Design requirements, 2005.

Ali, N., Fuchs, A., Neunaber, I., Peinke, J., and Cal, R. B.: Multi-scale/fractal processes in the wake of a wind turbine array boundary layer, J. Turbul., 20, 93–120, https://doi.org/10.1080/14685248.2019.1590584, 2019.

Anvari, M., Lohmann, G., Wächter, M., Milan, P., Lorenz, E., Heinemann, D., Tabar, M. R. R., and Peinke, J.: Short term fluctuations of wind and solar power systems, New J. Phys., 18, 063027, https://doi.org/10.1088/1367-2630/18/6/063027, 2016.

Boettcher, F., Renner, C., Waldl, H.-P., and Peinke, J.: On the statistics of wind gusts, Bound.-Lay. Meteorol., 108, 163–173, https://doi.org/10.1023/A:1023009722736, 2003.

Brokish, K. and Kirtley, J. L.: Pitfalls of modeling wind power using Markov chains, in: 2009 IEEE/PES Power Systems Conference and 320 Exposition, 1–6, https://doi.org/10.1109/PSCE.2009.4840265, 2009.

Carpinone, A., Langella, R., Testa, A., and Giorgio, M.: Very short-term probabilistic wind power forecasting based on Markov chain models, in: 2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, 107–112, https://doi.org/10.1109/PMAPS.2010.5528983, 2010.

Einstein, A.: Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, Ann. Phys., 322, 549–560, https://doi.org/10.1002/andp.19053220806, 1905.

Friedrich, R., Peinke, J., Sahimi, M., and Tabar, M. R. R.: Approaching complexity by stochastic methods: From biological systems to turbulence, Phys. Rep., 506, 87–162, https://doi.org/10.1016/j.physrep.2011.05.003, 2011.

Frisch, U.: Turbulence: The Legacy of A. N. Kolmogorov, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781139170666, 1995.

Hadjihoseini, A., Lind, P. G., Mori, N., Hoffmann, N. P., and Peinke, J.: Rogue waves and entropy consumption, EPL, 120, 30008, https://doi.org/10.1209/0295-5075/120/30008, 2018.

Hadjihosseini, A., Wächter, M., P.Hoffmann, N., and Peinke, J.: Capturing rogue waves by multi-point statistics, New J. Phys., 18, 013017, https://doi.org/10.1088/1367-2630/18/1/013017, 2016.

Haehne, H., Schottler, J., Waechter, M., Peinke, J., and Kamps, O.: The footprint of atmospheric turbulence in power grid frequency measurements, EPL, 121, 30001, https://doi.org/10.1209/0295-5075/121/30001, 2018.

Kolmogorov, A. N.: A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number, J. Fluid Mech., 13, 82–85, https://doi.org/10.1017/S0022112062000518, 1962.

Milan, P., Wächter, M., and Peinke, J.: Turbulent Character of Wind Energy, Phys. Rev. Lett., 110, 138701, https://doi.org/10.1103/PhysRevLett.110.138701, 2013.

Morales, A., Waechter, M., and Peinke, J.: Characterization of wind turbulence by higher-order statistics, Wind Energy, 15, 391–406, https://doi.org/10.1002/we.478, 2012.

Mücke, T., Kleinhans, D., and Peinke, J.: Atmospheric turbulence and its influence on the alternating loads on wind turbines, Wind Energy, 14, 301–316, https://doi.org/10.1002/we.422, 2011.

Nawroth, A. P. and Peinke, J.: Multiscale reconstruction of time series, Phys. Lett. A, 360, 234–237, https://doi.org/10.1016/j.physleta.2006.08.024, 2006.

Nawroth, A. P., Friedrich, R., and Peinke, J.: Multi-scale description and prediction of financial time series, New J. Phys., 12, 083021, https://doi.org/10.1088/1367-2630/12/8/083021, 2010.

Papaefthymiou, G. and Klockl, B.: MCMC for Wind Power Simulation, IEEE T. Energ. Convers., 23, 234–240, https://doi.org/10.1109/TEC.2007.914174, 2008.

Peinke, J., Tabar, M. R., and Wächter, M.: The Fokker–Planck Approach to Complex Spatiotemporal Disordered Systems, Annu. Rev. Conden. Ma. P., 10, 107–132, https://doi.org/10.1146/annurev-conmatphys-033117-054252, 2019.

Pesch, T., Schröders, S., Allelein, H. J., and Hake, J. F.: A new Markov-chain-related statistical approach for modelling synthetic wind power time series, New J. Phys., 17, 055001, https://doi.org/10.1088/1367-2630/17/5/055001, 2015.

Reinke, N., Fuchs, A., Nickelsen, D., and Peinke, J.: On universal features of the turbulent cascade in terms of non-equilibrium thermodynamics, J. Fluid Mech., 848, 117–153, https://doi.org/10.1017/jfm.2018.360, 2018.

Renner, C., Friedrich, R., and Peinke, J.: Experimental indications for Markov properties of small-scale turbulence, J. Fluid Mech., 433, 383–409, https://doi.org/10.1017/S0022112001003597, 2001.

Risken, H.: The Fokker-Planck-Equation: Methods of Solution and Application, vol. 2 of Springer series in synergetics, Berlin, Springer, 1996.

Siefert, M. and Peinke, J.: Joint multi-scale statistics of longitudinal and transversal increments in small-scale wake turbulence, J. Turbul., 7, N50, https://doi.org/10.1080/14685240600677673a, 2006.

Stresing, R. and Peinke, J.: Towards a stochastic multipoint description of turbulence, New J. Phys., 12, 103046, https://doi.org/10.1088/1367-2630/12/10/103046, 2010.

Wächter, M., Riess, F., Kantz, H., and Peinke, J.: Stochastic analysis of surface roughness, EPL, 64, 579–585, https://doi.org/10.1209/epl/i2003-00616-4, 2003.

Wächter, M., Milan, P., Mücke, T., and Peinke, J.: Power performance of wind energy converters characterized as stochastic process: applications of the Langevin power curve, Wind Energy, 14, 711–717, https://doi.org/10.1002/we.453, 2011.

Weber, J., Zachow, C., and Witthaut, D.: Modeling long correlation times using additive binary Markov chains: Applications to wind generation time series, Phys. Rev. E, 97, 032138, https://doi.org/10.1103/PhysRevE.97.032138, 2018.

WindEurope: Wind in power 2017, available at: https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2017.pdf (last accessed: 18 August 2020), 2018.