

Improved Ramping Behaviour Analysis model for scenario generation using Markov-chain

Sambeet Mishra, Esin Oren and Ivo Palu
Tallinn University of Technology, Tallinn, Estonia

December 28, 2018

Abstract

Power generation from wind farm varies with time. This power swing increases the challenges in full utilization of the resource. The power system planner need to ensure supply of power on demand. Accurate estimation of power production from non-dispatchable resources such as wind is essential for the planner to ensure the supply-demand balance and reserve power capacity. This paper investigates the power swings and introduces improved ramping behavior analysis model namely RBA_{θ} . The objective of the proposed model is to quantify the variations as events. Each event is described through parameters- Up-ramp, Down-ramp, Peak, Frequency and Persistence. For this investigation, wind power production data is obtained from a wind farm containing 8 turbines. The data is normalized and scaled between 0 and 1. The geographical location of are synthesized from coastal region of Estonia. In first step, the events are extracted from the time-series wind power production data. In step two, the rain-flow counting algorithm is used to identify the cycles in the time-series data. These cycles explain if a noticeable pattern occurs multiple times in the data. Subsequently in step 3, a spatial Markov chain is implemented that takes the location as input and produce a matrix of transition probabilities. Then in step 4, the sequence of events are predicted with reference to time. Subsequently in final step, the transition matrices containing spatial information and time-series events are combined to generate scenarios that are- most likely, least likely and probable. The scenarios are generated for a long term and short term period. A comparative analysis between temporal scenarios and spatio-temporal scenarios is presented to outline the importance of the proposed method.

Nomenclature

RBA Ramping Behaviour Analysis

1 Introduction

Wind energy is stochastic in nature due the flow of wind is a product of multiple natural phenomenon. When there is a small penetration of wind into the power systems, the uncertain behaviour of the wind power generation is treated as just another uncertainty on the demand side. Moreover, the conventional power stations covering this variability, require additional energy and reduce the environmental benefits. Forecasting is one of the many possible solutions to this problem, as well as an interconnected grid, energy storage technologies, demand-side management such as electric vehicles. Forecasting aims to model the uncertainties inherited by the grid through wind power production and thus is a necessary and cost-effective element for the optimal integration of wind power into energy systems. However, forecasting is never accurate and literature suggests providing bounds for the forecasts or confidence intervals. Wind power is intermittent by nature due to weather. A wind power ramp event can be termed as a sudden change in the output power over a set threshold. Mathematically, the absolute difference between power produced P_t in time t and $(t + \Delta t)$ that is above the set threshold \bar{P} is a ramp event as in (1). However the threshold value is subjective.

$$|P_{(t+\Delta t)} - P_t| > \bar{P} \quad (1)$$

The system operator (SO) has to keep the system balanced i.e, the generation must meet the demand at each point in time. Wind ramp events can be positive or negative based on the generation swings. If positive, then the wind turbine has to shut down to avoid accidents or damage to the system whereas if the swing is negative the SO has to find a replacement to mitigate the demand. From economical point of view, both the energy

not used and energy from an alternative resource are crucial. Each wind farm forecasts wind speed and power production from historical data over time with an objective to determine potential investment and operations. In long-term forecasting the events become insignificant due to time-stretch while in short-term forecasting of events are more accurate. Furthermore, the time interval Δt is typically 10 minutes for ramp events. The \bar{P} is either set to an absolute value for a wind park or a certain percentage of the generation depending on the installed capacity. The problem with this practice is that the peak generation capacity varies through seasons and turbine maintenance or new installations. Although the threshold is subjective to peculiarities of a wind park, the methods to classify ramp events is generic. This work draws its focus on the procedure to detect ramp events. In addition demonstrates the application on real data from wind park.

1.1 Previous work

There are many studies in the literature on characteristics of wind power, correlations of wind and wind power, forecasting wind and wind power output, scenario generation. In [1] the authors proposed a model to forecast ramp events as well. They modelled observed wind speeds into forecast models and converted this into power forecasts with the help of the power curve of the wind turbines. It suggests that the same method could be implemented for solar power plants. Another noteworthy one is a model-free forecast generation implemented with Generative Adversarial Networks (GAN). GANs have two de-convolutional, one that starts by generating random data and the out discriminates whether its input is coming from the generator or historical data. These two neural networks play a Nashville game while giving feedback to each other, both getting better over time until the generator generates data that is almost like a forecast so that the discriminator can not discriminate any more. [2] proposed a probabilistic forecasting method, utilizing a Neural Network(NN) to generate possible future scenarios, employing an objective function based on cumulative distribution functions and auto-correlation functions to train the NN, primarily teaching it their distribution. Again another [3] proposed a model to synthesized wind speed scenarios based on statistical parameters of wind and Markov chains. In contrast, [4] proposes a new heuristic to generate scenarios that use copulas instead of common correlation functions. [5,6] introduced the terminology for identification of ramp events, ramping behaviour analysis(RBA), which comprises the perspective used in this study. They also filtered and extracted events, and clustered them into groups. More studies have been performed on identifying ramp events in [2, 7, 8].

1.2 Contribution of this paper

This paper investigates the variations in the wind power generation and contributes to the variation quantification and prediction methodology. The proposed model can be used for operational planning of wind farms and injection of wind power into the system. RBA_θ is a novel method considering spatio-temporal information of power variations. The proposed model has the following advantages over the existing models:

Cleaning the data Prediction of discrete time-series data may contain some irregularities in form of noise. Therefore, pre-processing the data and removing the noise content is essential for accurate prediction.

Spatial information The location of the turbines has a significant impact of the production. Therefore, RBA_θ takes into account the spatio-temporal information: location of turbines and power production through a markovian transition matrix. This matrix has transition probabilities with likelihood from one state to another state.

Prediction of events The operational planning for a wind farm plans for responses to significant events such as high power generation and very low power generation. As a consequence, generating the scenarios for the events that might occur based on the historical pattern facilitates the planning.

2 Modelling Ramping Behaviour Analysis Algorithm

The wind power variations are classified by wind the ramp events. The RBA (Ramping Behaviour Analysis) is an algorithm that determines ramp events based on peak power. This algorithm has four components: rise time, fall time, ramp-up rate and ramp-down rate. The graphical representation in fig. 1 describes a ramp event and the associated components. The angle (θ) is with reference to the peak point for both up-ramp and down-ramp

events. The persistence describes the amount of time the peak event persists. The rise-time and fall-time are presented as Δt^{up} and Δt^{down} respectively.

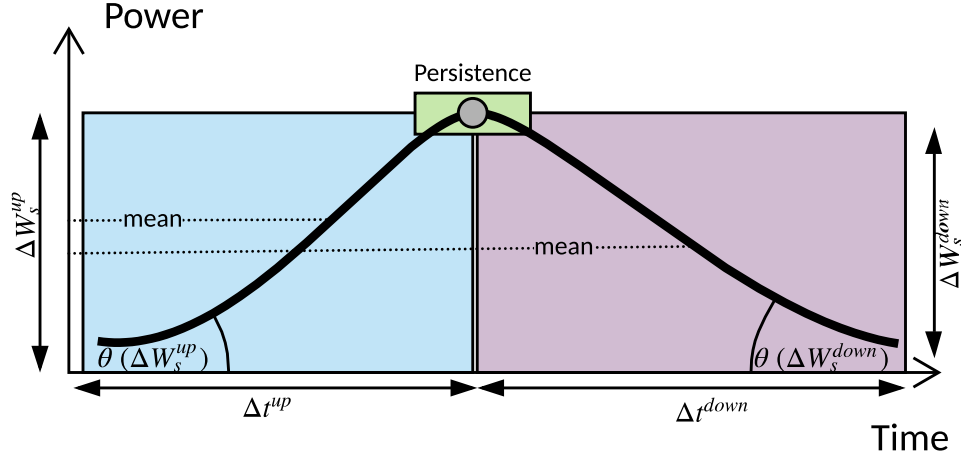


Figure 1: Wind ramp event classification

2.1 Wind power data

The data is obtained from the wind park with hourly time resolution. All data points coincide with each other date and time-wise for each turbine. Fig. 2 presents the power output from the wind park for 8760 hours. The data has certainly many power swings and it can be seen more clearly when a smaller time-window is chosen. However, the data is demonstrated with a top-view to show the pattern in seasons. For example, the winter season has dense power power production in comparison to lighter one in summer.

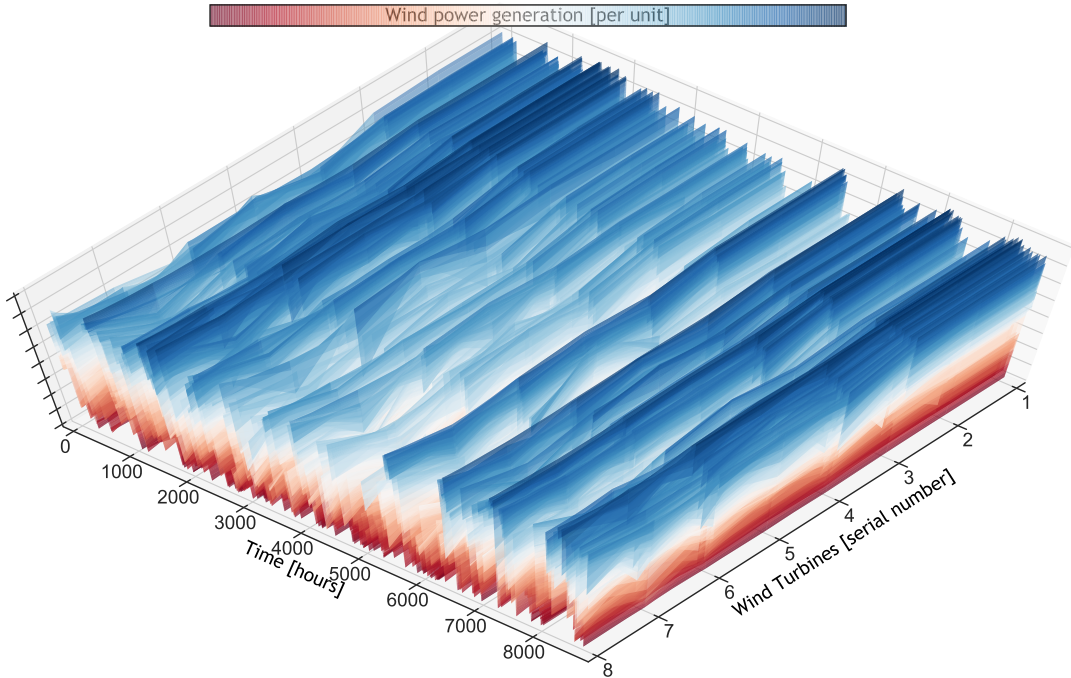


Figure 2: Wind power generation from Palidiski wind farm

wind production of a turbine is expressed through capacity factor: the ratio between the net power generation and the calibrated as in $\frac{P_{net}}{P_{rated}}$. The data are registered with a time stamp, however the relation between two subsequent observation is not. Again, the degree of noise content that is inherent to the data is substantial for

event detection. For example one outlier could either be a significant event. Thus raw data is smoothed through different smoothing techniques to capture the significant events.

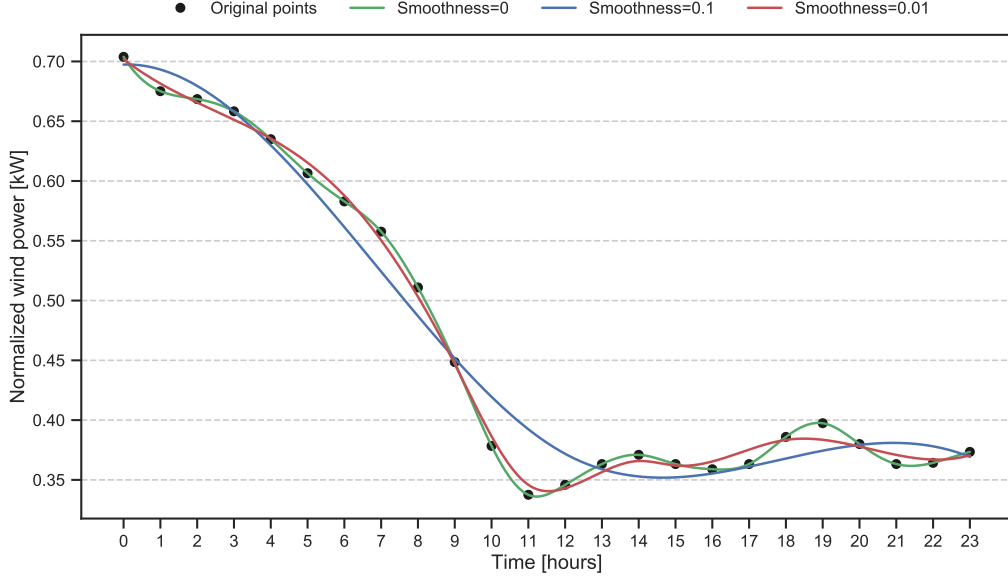


Figure 3: Wind power data smoothing using cubic spline smoothing technique

The wind power production data from the wind farm is pre-processed to smoothing technique. A cubic spline method is chosen that is a special case of spline interpolation. In (2) presents the S_i cubic spline function. Subsequently the coefficients are obtained starting from z_0 and then using the recurrence relation. This technique has smaller error than Newton polynomial and Lagrange polynomial [9]. In fig. 3 the original data and a comparison between smoothness 0, 0.1 and 0.01 for a 24 hour period is presented. It is evident that smoothness 0.01 is very close to the original trend in the wind power data.

$$S_i(x) = y_i + z_i(x - x_i) + \frac{z_{i+1} - z_i}{2 * (x_{i+1} - x_i)}(x - x_i)^2 \quad (2a)$$

$$z_{i+1} = -z_i + 2 * \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (2b)$$

A given function or signal can be transformed from time to frequency domain and other way round. The frequency domain transforms the linear differential to algebraic equations which are easy to solve. Furthermore, the latter provides qualitative behavior of the system: such as bandwidth, frequency response, gain, phase shift, power spectral density, eigenvalues to name a few. The focus on this thesis is limited to spectral density and FFT (Fast Fourier transformation). The Fourier transform of a function contains all the information about the original signal, and with this information, it is possible to reconstruct the function entirely by an inverse Fourier transform. This information includes amplitude and phase of each frequency present in the function. The Fourier transform of a discrete-time signal $x[n]$, $n = 0, \dots, N$ is called the discrete-time Fourier transform (DTFT), which provides a mathematical approximation of the full integral solution, and yields a periodic frequency spectrum. The DTFT of the sequence $x[n]$ denoted in (3a) is a function of a continuous frequency variable ω and $X(e^{j\omega})$ and is always periodic with period 2π . And (3b) represents the inverse DTFT of $x[n]$. DFT (Discrete Fourier Transform) can be obtained from the DTFT by evaluating through a discrete set of equally spaced frequencies [10].

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad (3a)$$

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n} d\omega \quad (3b)$$

A finite number of samples are selected to determine the spectrum. Then a window is generated by a multiplication of $x[n]$ by another sequence $w[n]$. Blackman window is selected for the study [11]. A time-domain representation of the same is presented in (4) where N is the *length* of Blackman window.

$$w(n) = 0.42 - 5 \cos\left(\frac{2\pi n}{N-1}\right) + 0.88 \cos\left(\frac{4\pi n}{N-1}\right) \quad \forall 0 \leq n \leq (M-1) \quad (4)$$

$$M = \begin{cases} N/2 & \text{if } N \text{ is even} \\ (N+1)/2 & \text{if } N \text{ is odd} \end{cases}$$

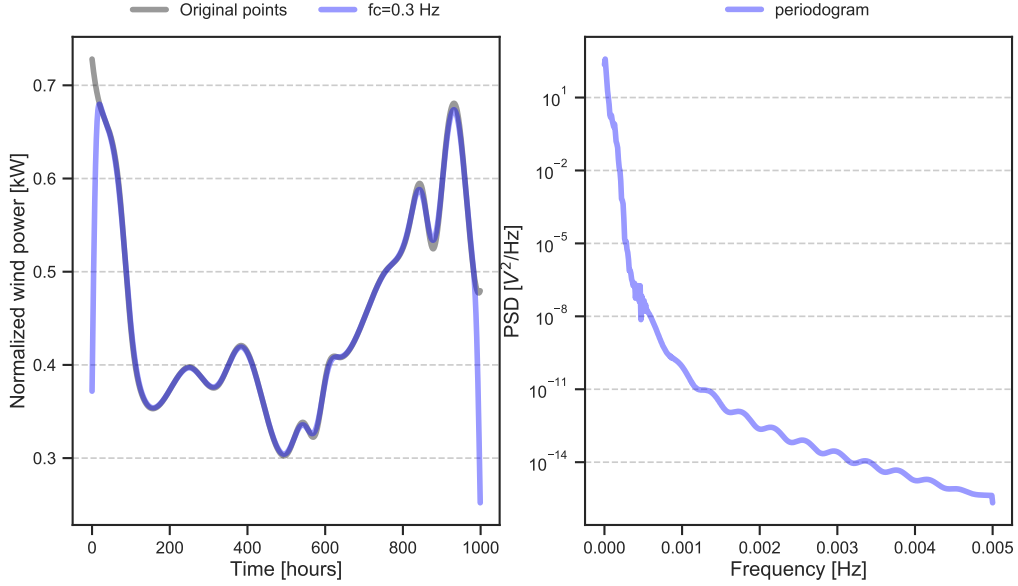


Figure 4: Frequency domain transformation: (A) filtered data with 0.3 Hz cut-off frequency (B) periodogram plot showing power spectral density

Fig. 4 presents the data in frequency domain after the transformation from time to frequency domain. A Blackman window filter is used to remove the noise content in the wind power data. A smaller time window is chosen to demonstrate the changes. In fig. 4(A) the original data and data with 0.3 Hz cut-off frequency (f_s) is presented. Power spectral density (Welch-periodogram) of $x[n]$ is presented in fig. 4(B). The data is pre-processed for removing the noise content in the data during measurement. For this, a smoothing technique is used to smooth the data. Following that, a time to frequency domain transformation is performed. In this process, the data is passed through Blackman filter. The subsequent section explains the RBA_θ and the extraction of events from the original time-series power production.

2.2 Improved Ramping Behaviour Algorithm

RBA (Ramping Behaviour Analysis) is a model to detect wind ramp events. Consider the wind power generation w_t at discrete time t and the consecutive measurement as w_{t+1} where the balance is $w_{\Delta t}$. A finite variation Δw therefore denoted as a ramp. Positive value of Δw becomes a positive ramp and otherwise is denoted as a negative ramp. Thus a ramp event Δw_s is defined as an event where a significant change in power generation takes place in a time period Δt . The significance is determined through the parameter T that stands for an adjustable threshold to neglect Δw values that are smaller than a set threshold T as in $\Delta w_s = \Delta w \mid \Delta w > T$. A detailed explanation of RBA is presented in the [5], [6]. The mathematical expression for identification of peak points is presented in (5a).

The angle between the interval and the change in amplitude is denoted as θ . It can be formulated as in (5). An algorithmic description of the RBA_θ is presented in algorithm 1. The description includes two functions. The first one extract the events and the second one evaluates the number of occurrences of those events. The algorithm takes the input of time varying wind data (vector), the total length of the time period (scalar). Then

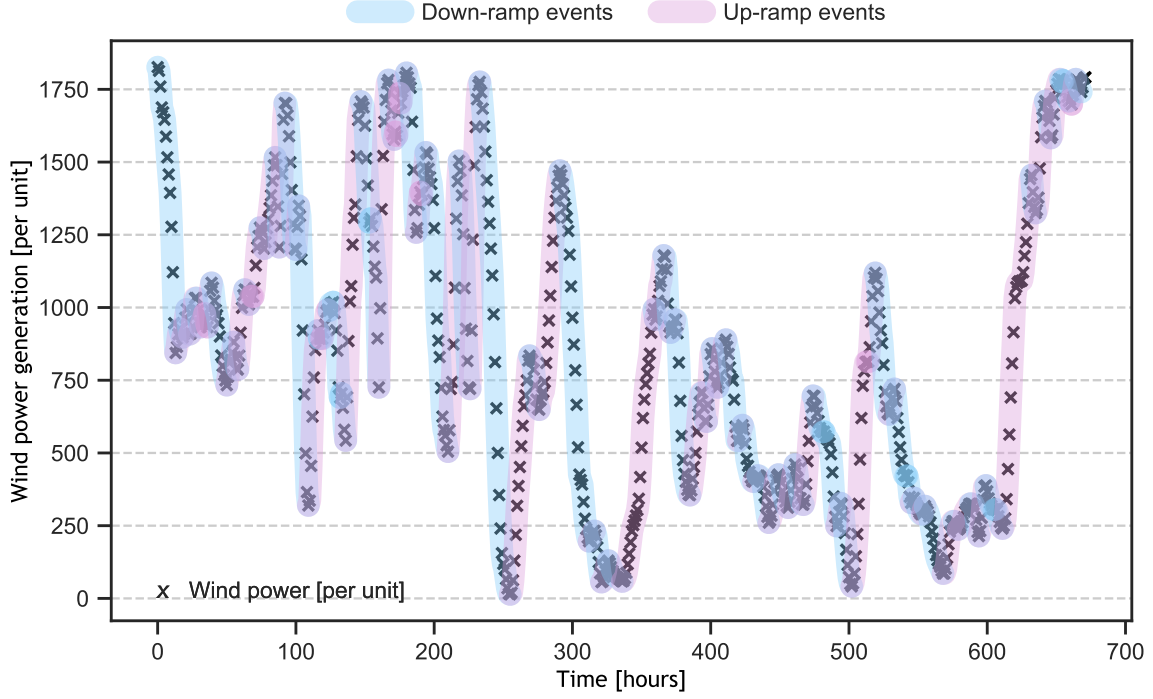


Figure 5: Wind ramp detection with 5% threshold

the events are identified with respect to peak points and event characteristics: ramp-up, ramp-down, rise-time, fall-time, angle of the peak with respect to adjacent minima and persistence values are the results. The angle θ between two significant changes Δw_s and Δt is presented in (5b).

$$\text{mean}(\Delta w_s) = \left(\frac{w_{s,t} + w_{s,(t+\Delta t)}}{2} \right) \quad (5a)$$

$$\theta(\Delta w_s) = \arctan(\Delta w_s, \Delta t) \quad (5b)$$

The threshold T was set to 0.01 or 5% of the nominal capacity of the wind turbine. Fig. 5 presents the resulting ramp event detection. The power production are marked with x marks and ramp-up and ramp-down events are highlighted through blue and purple. Arrows present the scheme used to classify the events, for instance a long red arrow describes a long ramp-up event wherein there are two peak points considered as one significant up event. Note that the first ramp up event persisted for longer period in comparison to the second and third. Similarly the third ramp event persisted for longer period than the other. Clearly persistence values explain the length of an event and thereby indicates the severity. In other words, in a specific period of time if there exists multiple peak points with or without slight difference then the multiple points are counted as one event persisting over that time period. Note that frequency and persistence are distinct from each other. While the frequency counts the number of occurrences of an event in the time-series data, the persistence values demonstrate how long an event has persisted with respect to time.

The frequency of occurrence is a measure to count these ramp events in the time horizon of the wind power data. Since this method identifies local peaks in the data with respect to the adjacent minimum point, the first ramp-up event ignores the first peak (at 0.16) and considers the peak at 0.20 in y-axis. Consequently the adjacent minimum point 0.13 is ignored and 0.12 is recorded. There is an overlap of the events at the meeting point this is because the event length includes both the starting and end point of the wind power data. In this way, events are aligned end-to-end. With that, RBA_θ identifies and counts the events while recording the event details at all stages. The table 1 presents the RBA_θ parameters for 10 sample events.

2.3 RBA_θ Frequency and Persistence

Frequency of an event can be defined as the number of times it occurs. Frequency of occurrences for Δw_s , Δt , $\text{mean}\Delta w_s$, $\alpha\Delta w_s$ with corresponding ranges are evaluated. The ranges or bins are assumed to be linearly spaced

Algorithm 1: RBA_θ algorithm pseudo code

Result: $w_{s,t}, w_{s,(t+\Delta t)}, (t + \Delta t), \Delta w_s, \theta(\Delta w_s), \text{mean}(\Delta w_s), \Delta w_s^f, \Delta t_s^f, (\theta(\Delta w_s))^f, \text{mean}(\Delta w_s^f)$

```

1 function  $RBA_\theta$  ;
  Input :  $T, w_t \in \mathbb{R}^+$ 
2  $\Delta w = w_{t+\Delta t} - w_t$  ;
3 for  $i \leftarrow 1$  to  $\text{length}(\Delta w)$  do
4   if  $\text{sign}(\Delta w[i]) == \text{sign}(\Delta w[i+1])$  then
5     concatenate( $\Delta w$ )
6   end
7 end
8 for  $i \leftarrow 1$  to  $\text{length}(\Delta w)$  do
9   if  $\Delta w > T$  then
10     $\Delta w_s \leftarrow \Delta w$ 
11  end
12 end
13 for  $i \leftarrow 1$  to  $\text{length}(\Delta w_s)$  do
14   if  $\text{sign}(\Delta w_s[i]) == \text{sign}(\Delta w_s[i+1])$  then
15     concatenate( $\Delta w_s$ )
16     events[i]  $\leftarrow w_{s,t}, w_{s,(t+\Delta t)}, t, (t + \Delta t), \Delta t_s, \theta(\Delta w_s), \text{mean}(\Delta w_s)$ 
17   end
18 end

```

Event	w_s^t	$\Delta w_s(t + \Delta t)$	t	$t + \Delta t$	Δw_s^t	Δt	$\theta \Delta w_s$	$\text{mean}(\Delta w_s)$
01	0.580	0.798	0.000	1.000	0.218	1.000	65.340	0.689
02	0.691	0.336	14.000	123.000	-0.355	109.000	-1.868	0.514
03	0.482	0.684	872.000	914.000	0.201	42.000	2.743	0.583
04	0.684	0.126	914.000	1083.000	-0.558	169.000	-1.891	0.405
05	0.541	0.126	1012.000	1083.000	-0.415	71.000	-3.347	0.333
06	0.126	0.367	1083.000	1150.000	0.241	67.000	2.063	0.246
07	0.215	0.683	1348.000	1466.000	0.468	118.000	2.270	0.449
08	0.683	0.515	1466.000	1523.000	-0.168	57.000	-1.685	0.599
09	0.522	0.291	1532.000	1589.000	-0.231	57.000	-2.320	0.407
10	0.291	0.714	1589.000	1658.000	0.422	69.000	3.503	0.502

Table 1: RBA_θ parameters

between 1 and 100. The algorithmic representation for the RBA_θ frequency function is presented in algorithm 2. The table 2 presents the start and end points for the event parameters for frequency evaluation. Input time series can be of different forms, for instance a triangular or step. For the former the persistence would be zero because there is only one peak for a time period. In case of the later there would be multiple points with or without slight variations with respect to a time period. Table 3 presents the parameters and corresponding the frequency of occurrences of events.

Algorithm 2: RBA_θ frequency function

Input : $\Delta w_s, \Delta t_s, (\theta(\Delta w_s)), \text{mean}(\Delta w_s)$
Result: $\Delta w_s^f, \Delta t_s^f, (\theta(\Delta w_s))^f, \text{mean}(\Delta w_s^f)$

```

1 for  $bin_1 = -1 \dots 1$  do
2   |  $w_{s,bin_1}$ 
3 end
4 for  $bin_2 = 1 \dots \max(\Delta t_s)$  do
5   |  $\Delta t_{s,bin_2}$ 
6 end
7 for  $bin_3 = -90 \dots 90$  do
8   |  $\theta(\Delta w_{s,bin_3})$ 
9 end
10 function frequency
    Input : param, bin
11 for  $i \leftarrow \text{to length}(bin)$  do
12   | for  $k \leftarrow \text{to length}(param)$  do
13     | if  $param[k] \leq bin[k+1]$  then
14       |  $f[i] += 1$ 
15     | end
16   | end
17 end
18 for  $i \leftarrow \text{to length}(bins)-1$  do
19   | for  $k \leftarrow \text{to length}(param)$  do
20     |  $frequency[i] = f[i]$ 
21   | end
22 end
23  $\Delta w_s^f = \text{frequency}(\Delta w_s, \Delta w_{s,bin_1})$ 
24  $\Delta t_s^f = \text{frequency}(\Delta t_s, \Delta t_{s,bin_2})$ 
25  $\theta(\Delta w_s^f) = \text{frequency}(\theta(\Delta w_s, \Delta w_{s,bin_3}))$ 
26  $\text{mean}(\Delta w_s^f) = \text{frequency}(\text{mean}(\Delta w_s, \Delta w_{s,bin_4}))$ 

```

Event parameters	Starting point	Ending point
Δw_s	-1	01
Δt	01	$\max(\Delta t)$
$\text{mean}(\Delta w_s)$	$\min(\text{mean}(\Delta w_s))$	$\max(\text{mean}(\Delta w_s))$
$\alpha(\Delta w_s)$	-90°	90°

Table 2: Frequency measures of events with start and end points

The persistence can be defined as the act of continuing to exist past the usual time. Therefore the persistence event is a series of points that stays within the defined threshold value for longer than a predefined time limit. The algorithmic representation of persistence is presented in algorithm 3. The time period is chosen to be two days. In this period all the events extracted are sorted and the first event that has persisted for highest time is saved as a persistent event. The table 4 presents the persistence values for ten events. The persistence values for $w_{s,t}$ and $w_{s,t+\Delta t}$ are presented through time t , time difference $t + \Delta t$, mean and time difference Δt .

Event	Δw_s	Δw_s^f	Δt	Δt^f	$\theta \Delta w_s$	$\theta \Delta w_s^f$	mean(Δw_s)	mean(Δw_s^f)
01	0.218	14	1	1	65.340	1	0.689	5
02	-0.355	10	109	9	-1.868	167	0.514	5
03	0.201	20	42	13	2.743	37	0.583	3
04	-0.558	7	169	5	-1.891	167	0.405	9
05	-0.415	4	71	16	-3.347	23	0.333	9
06	0.241	18	67	17	2.063	159	0.246	4
07	0.468	4	118	8	2.270	159	0.449	4
08	-0.168	26	57	15	-1.685	167	0.599	2
09	-0.231	16	57	15	-2.320	167	0.407	9
10	0.422	9	69	12	3.503	37	0.502	3

Table 3: RBA_θ frequency of events

Algorithm 3: RBA_θ persistence function

Input : $mean, \Delta t$
Result: $\overline{mean}, \overline{\Delta t}$
1 periods = $t + 1$
2 bins = linspace(0, t , periods)
3 $t = \text{length}(\text{filtered})$
4 **for** i in range($\text{length}(\text{bins}-1$)) **do**
5 **for** k in range($\text{length}(mean, \Delta t)$) **do**
6 **if** $\text{bins}_i < mean_k, \Delta t_k < \text{bins}_{i+1}$ **then**
7 sort ($\overline{mean}, \overline{\Delta t}$)
8 **end**
9 **end**
10 **end**

Event	$w_{s,t}$	$w_{s,t+\Delta t}$	t	$t + \Delta t$	mean	Δt
01	0.453	0.302	100	477	0.378	377
02	0.439	0.591	102	832	0.515	730
03	0.265	0.416	1056	1389	0.340	333
04	0.583	0.411	1624	1999	0.497	375
05	0.452	0.302	100	477	0.377	377
06	0.438	0.590	102	831	0.514	729
07	0.276	0.428	1054	1392	0.352	338
08	0.571	0.408	1622	1999	0.489	377
09	0.482	0.331	94	466	0.406	372
10	0.471	0.621	96	889	0.546	793

Table 4: RBA_θ persistence of events

2.4 Rainflow Counting Cycle for Event extraction

Rainflow counting algorithm [12] was developed to be used in the analysis of fatigue data in order to reduce a spectrum of varying stress into a set of simple stress reversals. The input to the algorithm is a simple series of peaks and valleys i.e., local maxima and minima, that form hysteresis loops. Closed loops are full cycles, and open loops are half cycles. The algorithm uses a change in slope as an indicator that the time series is going through a peak or valley. Only the magnitude of the peak or valley is then entered into the Rainflow counting algorithm.

Events	w_s^t	Cycles
01	0.218	0.5
02	-0.355	0.5
03	0.201	1
04	-0.558	1
05	-0.415	0.5
06	0.241	0.5
07	0.468	0.5
08	-0.168	1
09	-0.231	1
10	0.422	0.5

Table 5: Rainflow counting cycle for first ten events

It was implemented as presented in [13]. This algorithm is used to extract the cycles, with modifications to extract the starting and ending points of ramp events hence the time range, the starting and ending power of the events hence the amplitude, the angle of the event and the cycle with some minor modifications. A sample parameter values for the rainflow counting cycle is presented in the table 5. The full cycles are presented as 1 and half-cycles as 0.5. For example, the events 3 and 4 have full cycles.

3 Time-series Prediction using Deep Learning

Recurrent neural networks (RNNs) might be less potent than feed-forward networks but they are applicable in some cases which require communication between each neuron as it is in our task.

LSTM is a specified RNN architecture, and it is one of the most applicable networks for forecast time series. RNN is possibly the closest network that mimics the human brain. As it can be understood from the term “long-short”, the aim of this network is to model temporal sequences and their long-term dependencies. LSTM uses stochastic gradient descent as an optimization tool. The reason behind RNN outperforms than feed-forward networks is it contains cyclic connections. Thanks to these connections it is easier to model sequence data [14].

4 Spatial Markov Chain applications

The theory of spatial dynamics states that the power production from individual turbine can be quite different even though they are physically located in one farm and quite close to each other. The difference might arise from wake effect, terrain condition and other environmental effects. To capture this relation, a spatial dynamics approach has been considered. Specifically, spatial Markov chain model [15, 16]. In this section, a finite state spatial dynamics model is presented that is used to derive the transition matrix for the wind turbines in a wind farm. A Markov chain has a series of states n which are mutually exclusive of each other thus applicable to discrete data-set. A transition matrix has transition probabilities from one state to another. The fig. 6 presents the spacial relationship between wind turbines. The location information is presented in the table 6. The polygons are created using the GeoDa [17] program.

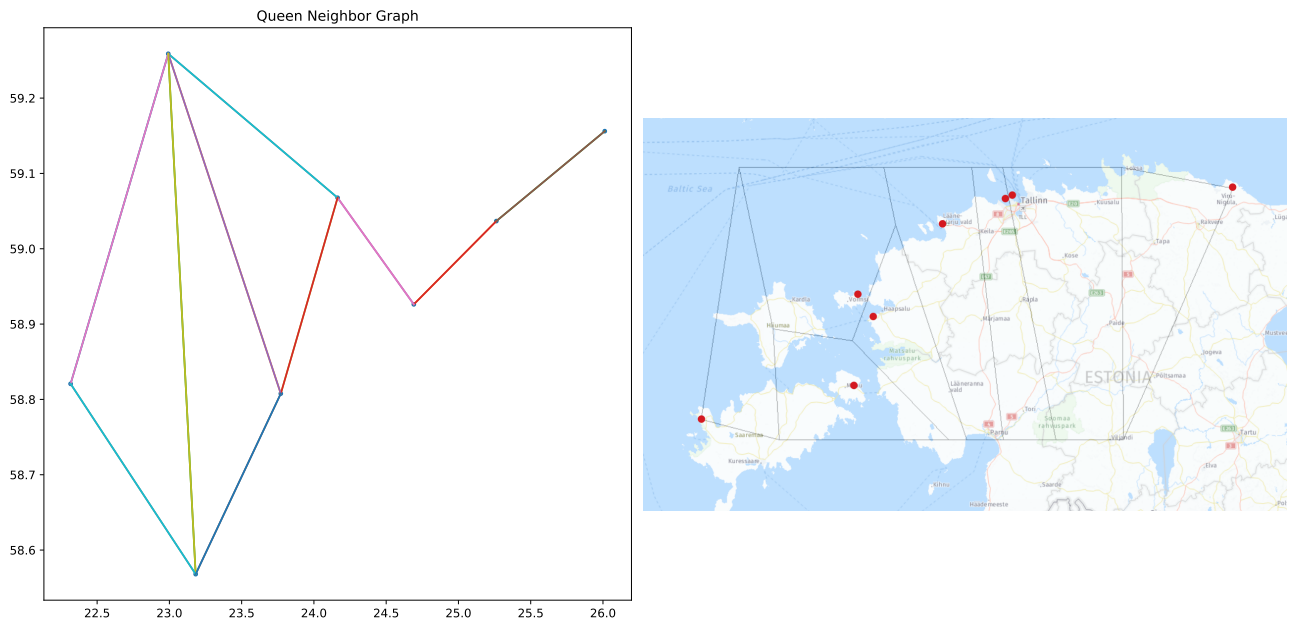


Figure 6: Spatial relation between wind turbine locations: on left queens neighbor graph and on right the actual locations in Estonia

Wind turbine	Longitude	Lattitude	Location
01	59.4784	24.6923	Paljassaare conservation area,Tallinn
02	59.4646	24.6567	Kopli peninsula,Tallinn
03	59.3416	24.1315	Paldiski, Harju county
04	59.0257	23.3329	Norrby, Lääne county
05	58.4434	21.9644	Kõruse, Saare county
06	58.9245	23.4778	Rohuküla, Lääne county
07	59.4953	26.6583	Iila, Lääne-Viru county
08	58.5955	23.2965	Mõega, Saare county

Table 6: Location of wind turbines

4.1 Markov process

A deterministic Markov chain has the following property

$$P(X_{n+1}^t \in A | X_0^t \dots = x_n) = P(X_{n+1}^t \in A | X_n^t = x_n) \quad (6)$$

In 6 the Markov property of the future state $n + 1$ depends only on the current state n not the past state $n - 1$. The process spread over all time $t_0 \leq t_1 \leq \dots t_n \leq t_{n+1}$ and state space $x_0, \dots x_n$. The transition kernel for the probability matrix can be expressed as in 7.

$$P_{s,t}(x, A) := P(x_t \in A | X_s = x) \quad (7)$$

The kernel generates transition probability value A that the process takes at time t given that at time s the value was x . The transition matrix is expressed as $P(p(x, y))_{x, y \in S}$. The initial state is $P(X_0 = x)_{x \in S}$. The distribution at time $t = 1$ is presented in (9). One step transition probability is expressed in (10).

$$P(X_1 = x) = \sum_{x_0 \in S} P(X_0 = x_0) * P(X_1 = x | X_0 = x_0) \quad (8)$$

$$p(x, y) := P(X_{n+1} = y | X_n = x) \quad (9)$$

$$(10)$$

5 Scenario generation through spatial Markov transition matrices

- why markov transition matrix to make scenarios
- long and short term scenario
- 4*4 matrix with 2 plots for short and 2 for long term scenarios with original data

6 Conclusion

Here is the conclusion and future scope of $RBA_\theta \dots$

References

- [1] Laura Bianco, Irina V Djalalova, James M Wilczak, Joel Cline, Stan Calvert, Elena Konopleva-Akish, Cathy Finley, and Jeffrey Freedman. A wind energy ramp tool and metric for measuring the skill of numerical weather prediction models. *Weather and Forecasting*, 31(4):1137–1156, 2016.
- [2] Mingjian Cui, Jie Zhang, Anthony R Florita, Bri-Mathias Hodge, Deping Ke, and Yuanzhang Sun. An optimized swinging door algorithm for identifying wind ramping events. *IEEE Transactions on Sustainable Energy*, 7(1):150–162, 2016.
- [3] Selin Karatepe and Kenneth W Corscadden. Wind speed estimation: incorporating seasonal data using markov chain models. *ISRN Renewable Energy*, 2013, 2013.
- [4] Michal Kaut. A copula-based heuristic for scenario generation. *Computational Management Science*, 11(4):503–516, 2014.
- [5] S. Mishra, M. Leinakse, and I. Palu. Wind power variation identification using ramping behavior analysis. *Energy Procedia*, 141:565–571, 2017.
- [6] S. Mishra, M. Leinakse, I. Palu, and J. Kilter. Ramping behaviour analysis of wind farms. *EEEIC2018*, 1(1):1–7, 2018.

- [7] Arthur Bossavy, Robin Girard, and George Kariniotakis. Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy*, 16(1):51–63, 2013.
- [8] Arthur Bossavy, Robin Girard, and Georges Kariniotakis. A novel methodology for comparison of different wind power ramp characterization approaches. In *EWEA 2013-European Wind Energy Association annual event*, pages 6–pages, 2013.
- [9] Carl De Boor, Carl De Boor, Etats-Unis Mathématicien, Carl De Boor, and Carl De Boor. *A practical guide to splines*, volume 27. Springer-Verlag New York, 1978.
- [10] James H McClellan, Ronald W Schafer, and Mark A Yoder. *Signal processing first*. Pearson education Upper Saddle River, NJ, 2003.
- [11] Prabhakar Agarwal, S Pratap Singh, and V kumar Pandey. Mathematical analysis of blackman window function in fractional fourier transform domain. In *Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 2014 International Conference on*, pages 120–125. IEEE, 2014.
- [12] Stephen D Downing and DF Socie. Simple rainflow counting algorithms. *International journal of fatigue*, 4(1):31–40, 1982.
- [13] Jennifer M Rinker and Henri P Gavin. Including sn-curve uncertainty in fatigue reliability analyses of wind turbines. In *Model Validation and Uncertainty Quantification, Volume 3*, pages 375–381. Springer, 2014.
- [14] Haşim Sak, Andrew Senior, and Françoise Beaufays. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *Fifteenth annual conference of the international speech communication association*, 2014.
- [15] Sergio J Rey. Spatial empirics for economic growth and convergence. *Geographical analysis*, 33(3):195–214, 2001.
- [16] Steven F Carle and Graham E Fogg. Modeling spatial variability with one and multidimensional continuous-lag markov chains. *Mathematical Geology*, 29(7):891–918, 1997.
- [17] Luc Anselin, Ibnu Syabri, and Youngihn Kho. Geoda: an introduction to spatial data analysis. *Geographical analysis*, 38(1):5–22, 2006.