An edge model for the evaluation of wind power ramps characterization approaches

Arthur Bossavy, Robin Girard, Georges Kariniotakis

An edge model for the evaluation of wind power ramps characterization approaches

Arthur Bossavy[1], Robin Girard[2], and George Kariniotakis[3]

[1,2,3]MINES ParisTech, PSL - Research University,, PERSEE - Center for Processes, Renewable Energies & Energy Systems,, CS 10207 rue Claude Daunesse, 06904 Sophia Antipolis Cedex, France.

Abstract

Wind power forecasting is a recognized means of facilitating large-scale wind power integration into power systems. Recently, there has been focus on developing dedicated short-term forecasting approaches for large and sharp wind power variations, so-called "ramps". Accurate forecasts of specific ramp characteristics (e.g., timing, probability of occurrence, etc.,) are important, as related forecast errors may lead to potentially large power imbalances, with a high impact on the power system. Various work about ramps' periodicity or predictability has led to the development of new characterization approaches. However, a thorough analysis of these approaches has not yet been carried out. Such an analysis is necessary to ensure the reliability of subsequent conclusions on ramps' characteristics. In this paper, we propose a comprehensive framework for evaluating and comparing different characterization approaches of wind power ramps. As a first step, we introduce a theoretical model of a ramp inspired from edge-detection literature. The proposed model incorporates some important aspects of the wind-power production process so as to reflect its non-stationary and bounded aspects, as well as the random nature of ramp occurrences. Then, we introduce adequate evaluation criteria from signal-processing and statistical literature, in order to assess the ability of an approach for reliably estimating ramp characteristics (i.e., timing, intensity). Based on simulations from this model and using the evaluation criteria, we study the performance of different ramp detection filters and multi-scale characterization approaches. Our results show that some practical choices in wind-energy literature are inappropriate, while others, namely from signal-processing literature, are preferable.

1 Introduction

Wind power forecasting is recognized as a means of facilitating large-scale wind power integration into power systems. Considerable R&D in the last 25 years has resulted in the development of numerous approaches. For a literature overview, we refer to [1, 2]. At an operational level, the applied models are generally accurate. Shortcomings often relate to challenging and extreme situations, but also to specialized forecasts for various business processes and their integration into decision-making tools.

Often the aim of ramp forecasting approaches is improving forecasts of specific ramp characteristics, like the ramps' timing or probability of occurrence. In related literature however, we observe that there is no standard definition of what a ramp is. Various work about ramps' periodicity or predictability has led to the development of new characterization approaches [3, 4, 5]. However, a thorough analysis of these approaches, which is necessary to ensure the reliability of subsequent conclusions on ramps characteristics, has not yet been done.

Characterizing a ramp in a formal way is not straightforward. Firstly, it requires consensus on a rigorous definition of the event, which may not yet exist. Understanding the pros and cons of different ramp characterization approaches in wind-energy literature is necessary and requires a dedicated mathematical model of a ramp. In edge-detection literature, the development and evaluation of a detection approach generally involve an edge[1] model and associated evaluation criteria (e.g., Canny's criteria [8]). There is no such paradigm in wind-energy literature. In [3] for instance, a singular form of filtering approach is reported. This approach does not specify a priori the nature of variations to be detected, nor that of the noise wherein the detection takes place. We do not know what geometry of variation, if there is one, the approach can detect, nor what the latter behavior could be with respect to other elements of a wind-power production signal (e.g., noise, nearby edges, etc.,). In [4], the authors propose detecting edges in a wind-power signal from the local maximum in the absolute response of a derivative filter. They used a Prewitt filter [9], which has proved efficient in some theoretical work but not optimal, for detecting edges shaped like a step [8]. Further investigations should determine if the use of such a filter is appropriate with respect to hypotheses, which better reflect the characteristics of a wind-power production signal.

Beyond the choice of a particular filter, a key issue in the detection of edges in a signal is the selection of one (or more) appropriate scale(s). Through a scale(s) selection process, the goal is usually to optimize a compromise between noise reduction and limitation of interactions due to the presence of neighboring edges [10]. Through a multi-scale detection approach, the goal may be to define an absolute characterization of an edge, thus unrelated to a particular scale. This characterization may rely on the estimation of some parameters defining the geometry of a particular edge model [11, 12]. In literature related to wind energy, a characterization of ramps based on a multi-scale approach has been recently proposed in [5]. This characterization relies on the definition of a measure of variations from the sum of filter responses at different scales. However, the proposed approach's analysis is restricted to the study of some sample situations. In addition, the approach's proposal raises interest in alternative approaches.

In this paper, we propose a comprehensive framework for evaluating and comparing different characterization approaches of wind power ramps. As a first step, we introduce a theoretical model of a ramp inspired from edge-detection literature. The proposed model incorporates some important aspects of the wind-power production process to reflect its non-stationary and bounded aspects, as well as the random nature of ramp occurrences. Then, adequate evaluation criteria from signal-processing and statistical literature are introduced to assess an approach's ability to reliably estimate ramp characteristics (i.e., timing, intensity). Based on simulations from this model and using the evaluation criteria, we study the performance of different ramp-detection filters and multi-scale characterization approaches. The conclusions illustrate the value of each approach: they show that some filters and approaches generally used in the wind-energy community for ramps' characterization and forecasting are inappropriate, while some approaches from edge-detection literature are preferable.

The paper is organized as follows: in Section 2 we describe the proposed ramp model. First, we define a ramp profile, where the randomness in the occurrence of two consecutive ramps is modeled. To this profile we add some noise so as to reproduce the fastest fluctuations of wind power production. Through modeling of the latter, we strive to adequately represent the bounded aspect of the wind-power production process, as well as some of its non-stationary aspects. In this section we also define the performance criteria, which together with the proposed ramp model, constitute our evaluation framework. In Section 3, we study the performance of state-of-the-art filters in detecting and localizing ramps. In Section 4, we focus on the performance of different scale selection and combination approaches. Finally, some summary conclusions are given in Section 5.

2   A framework for evaluating ramps' detection approaches

In this section, we describe the proposed framework for the evaluation of ramps' detection approaches. First, we define a theoretical edge model suitable for the particular case of wind power ramps' representa-

---

[1]In edge-detection literature, the term "edge" is commonly used to denote a sharp and large level change in a signal. The term "ramp" then generally refers to a particular geometry of edge (see [6] or [7] for instance). From now on, we will use the terms "edge" and "ramp" in an undifferentiated manner.

tion. We model the evolution of time separating edges in a signal, since it is a key parameter influencing detection performance [10]. Then, we describe the parameters' values used to simulate wind power ramps from the proposed model. Using a detection approach on a signal with simulated variations allows the ramps' detection results to be objectively counted. Finally, we introduce ramp detection as a supervised classification exercise and derive appropriate evaluation criteria.

2.1 Defining a theoretical model of ramps

In signal-processing literature [8, 13, 14, 15], an edge model is usually based on a decomposition of a signal p_t of the following form:

$$p_t = R_t + \epsilon_t, \tag{1}$$

where R_t denotes an edge profile and ϵ_t a noise. The definition of R_t enables the detection of a particular geometry of edge [6, 7, 8]. Its most important detection parameters are: the amplitude of an edge, its duration (or related, e.g., the edge blur), and the duration between two consecutive edges [13, 14]. In the latter, the duration is assumed to be constant. However, the random nature of variations in a non-stationary process - such as ramps in wind-power production - cannot be adequately represented this way.

In the proposed ramp model, we consider the duration between two consecutive ramps, as well as the duration of ramps itself as random variables. The ramp profile R_t represents a succession of constant production episodes joined by segments, namely the ramps. A ramp is then completely defined by associating a duration T_i with an amplitude A. The amplitude of ramps in wind-power production also has a random nature. Studying it requires an approach to estimate the amplitude of edges in a signal (see [11], [12] or [16], for instance). To limit the size of our study, we chose to study the performance of detection approaches from signals where ramps are of equal amplitude. Considering different amplitude values from different simulations should allow, to some extent, the study of the influence of this parameter on detection performances.

Considering null production episodes alternating with constant production episodes of value A, R_t is the succession of N *elementary profiles* R^k, each of random duration:

$$R = (R^1, \ldots, R^k, \ldots, R^N), \tag{2}$$
$$R^k = (R_1^k, \ldots, R_t^k, \ldots, R_{T^k}^k), \ 1 \leq k \leq N,$$
$$T^k = T_1^k + T_2^k + T_3^k + T_4^k,$$

$$R_t^k = \begin{cases} 0 & \text{if } t \in [0, T_1^k], \\ \frac{A}{T_2^k}(t - T_1^k) & \text{if } t \in [T_1^k, T_1^k + T_2^k], \\ A & \text{if } t \in [T_1^k + T_2^k, T_1^k + T_2^k + T_3^k], \\ -\frac{A}{T_4^k}(t - T^k) & \text{if } t \in [T_1^k + T_2^k + T_3^k, T^k], \end{cases} \tag{3}$$

where T_1^k (resp. T_3^k) denotes the duration of a null (resp. equal to A) production episode, and T_2^k (resp. T_4^k) the duration of an increasing (resp. decreasing) ramp. A representation of the elementary profile R^k on which the proposed ramp model is built, is shown in Figure 1. The simulation of a complete ramp profile R then comes from the sampling T_i^k, $k = 1, \ldots, N$, $i = 1, \ldots, 4$, where T_i^k are assumed to be independent truncated exponential variables with support $[\theta_m^i, \theta_M^i]$, and expectation λ_{T_i}.

To achieve realistic assumptions, we did an analysis based on power measurements sampled every 10 minutes and covering more than 18 months of production from three French wind farms. In Figure 2 (graph on the left) is shown the distribution of the signal production derivative of a wind farm estimated from the Prewitt filter (see [4]). We can see in this figure that the distribution of the power production derivative is more or less symmetrical (see also [17]). While we observed more decreasing than increasing variations among variations of lower intensity, increasing and decreasing variations are approximately
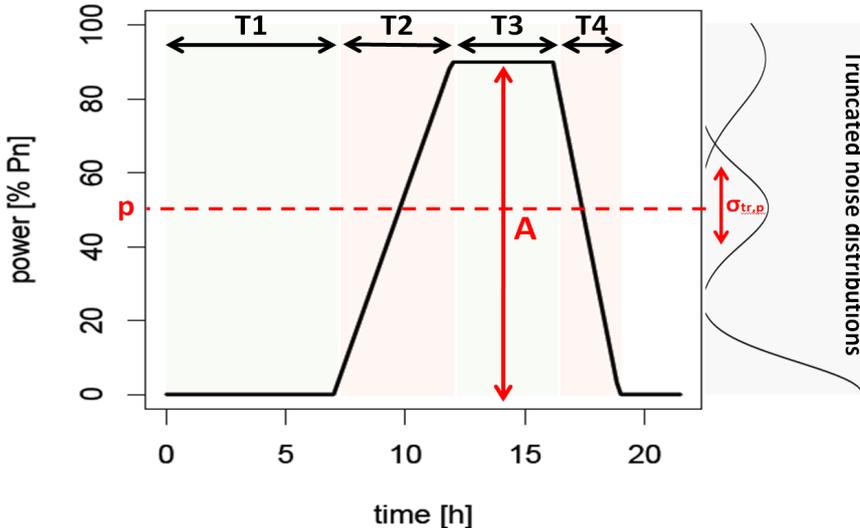
3

Figure 1: **Elementary ramp profile and noise distribution** – Diagram showing an elementary ramp profile R^k defined by Equation (3). Represented here are: the amplitude A of a ramp, the duration T_i, $i = 1, \ldots, 4$ of ramps and periods between consecutive ramps, and the distribution of the noise associated to a production level p in R^k. Power level is expressed in percentage of the nominal power P_n.

the same frequency for variations of higher intensity (see Figure 2, graph on the right). For ramps that are assumed to be of equal amplitude, this translates to an identical sharpness between increasing and decreasing ramps. In our model, we then assume that T_2 and T_4 follow the same probability law, i.e., $\theta_m^2 = \theta_m^4$, $\theta_M^2 = \theta_M^4$ and $\lambda_{T_2} = \lambda_{T_4}$. As low production is more common than high production, it suggests that episodes with low production values are generally longer than those with high production values. We will then assume that $\lambda_{T_1} > \lambda_{T_3}$.

To reproduce the fastest fluctuations of wind power production, we consider a white noise with truncated Gaussian distribution of support $[-p, 100 - p]$:

$$\epsilon_t \sim \mathcal{N}_{-p}^{100-p}(\mu_{\mathrm{tr},p}, \sigma_{\mathrm{tr},p}^2), \tag{4}$$

where p denotes the power production level represented by the ramp profile R_t at instant t (in percentage of the nominal capacity P_n of the wind farm), $\mu_{\mathrm{tr},p}$ the mean and $\sigma_{\mathrm{tr},p}$ the associated production standard deviation. Note that since noise distribution is derived from a centered normal distribution whose truncation generally is not symmetric (except for $p = 50\%P_n$), its mean $\mu_{\mathrm{tr},p}$ generally does not equal 0. Actually, $\mu_{\mathrm{tr},p} > 0$ when $p < 50\%P_n$ and $\mu_{\mathrm{tr},p} < 0$ when $p > 50\%P_n$. Note also that $\sigma_{\mathrm{tr},p} < \sigma_p$, where σ_p denotes the standard deviation of the initial (i.e. untruncated) normal distribution. Formulae defining $\mu_{\mathrm{tr},p}$ and $\sigma_{\mathrm{tr},p}$, depending on the untruncated normal distribution's first two moments, can be found in [18] (Chapter 3, p. 59).

The sigmoid form of the wind turbines' power curve causes wind variability to be amplified during the wind-to-power conversion process at intermediate wind speeds, while lowered at low and high wind speeds. Then, wind power variability first increases before stagnating, and ultimately decreases with wind power level (see Figure 3). However, data analysis suggested that a piecewise linear increase keeping power variability constant at higher power levels, even at those close to nominal power, was an adequate modelling of the production standard deviation:

$$\sigma_{\mathrm{tr},p}(p) = a_1 + a_2 p - a_2(p - p_1)_+, \tag{5}$$

where $a_i > 0, \forall i$, $(.)_+$ denotes the positive part of a real number, and $p_1 \in]0, 100[$ denotes the power level above which the production standard deviation remains at a constant maximal value.
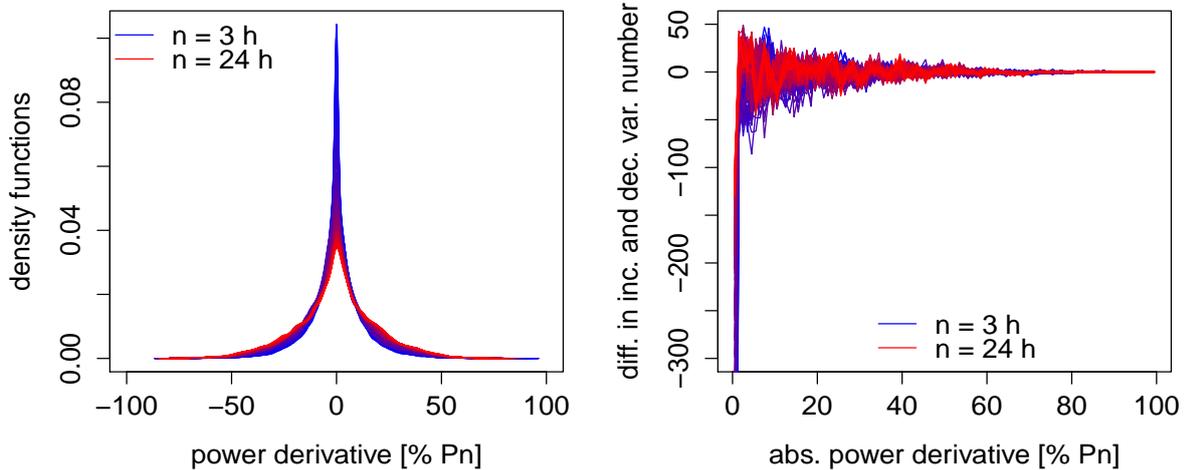
Figure 2: **Distribution of the power production derivative** – On the left: a kernel density estimation of the distribution of a wind-power production signal derivative. The latter has been estimated using a Prewitt filter (see [4]), for filter widths ranging from $n = 3$ h to $n = 24$ h, with increments of 20 minutes. On the right: the difference between the number of increasing and decreasing variations, depending on their intensity. While decreasing variations outnumber increasing variations among variations of lower intensity, they have approximately the same frequency for variations of higher intensity.

Some alternatives may represent wind-power distribution more appropriately. For short-term production forecasting, production distribution is sometimes identified (through centering and/or scaling transformations) as a beta-distribution [19] or as a logit-normal distribution incorporating [20], or not [21], the presence of punctual masses at the null and maximal production values. Here, a more refined modeling of wind-power distribution could offer further insight into the performance of ramp-detection approaches. While additional efforts in this direction provide valuable work perspectives, the latter remain outside the scope of this study.

2.2 Experimental conditions and simulations

The way noise distorts the geometry of a ramp depends on the production level at its beginning and end. The behavior of a detection approach may then vary with the considered production levels as well. However, our study is limited to the detection of ramps with null minimum production level, just as given by Equations (2) and (3). The ramp profile we are considering is then suitable for the representation of (not too much aggregated) production with recurrent null production episodes. We will consider ramps of various amplitudes: $A = 50, 60, 70, 80, 90$ and $100\% P_n$.

The duration of ramps and in-between consecutive ramps are key parameters for analyzing a detection approach's performance. We analyze such a performance by making these parameters vary, then by simulating different production conditions. The average duration of respectively null and constant high (equal to A) production episodes are assumed to be related by a multiplicative factor c, i.e. $\lambda_{T_1} = c\,\lambda_{T_3}$, $c \geq 1$. We consider values of $c = 1, 1.5, 2, 3, 5$, and $\lambda_{T_1} = 6, 12, 24, 48, 96$ h. Such values may not represent any real ramp occurrence frequency. Determining realistic values is precisely the purpose of estimation approaches we need to evaluate first. Anyway, for such values to be relevant here, the most important aspect is to consider them in conjunction with chosen detection scales, in order to reveal the characteristic behavior of detection approaches.
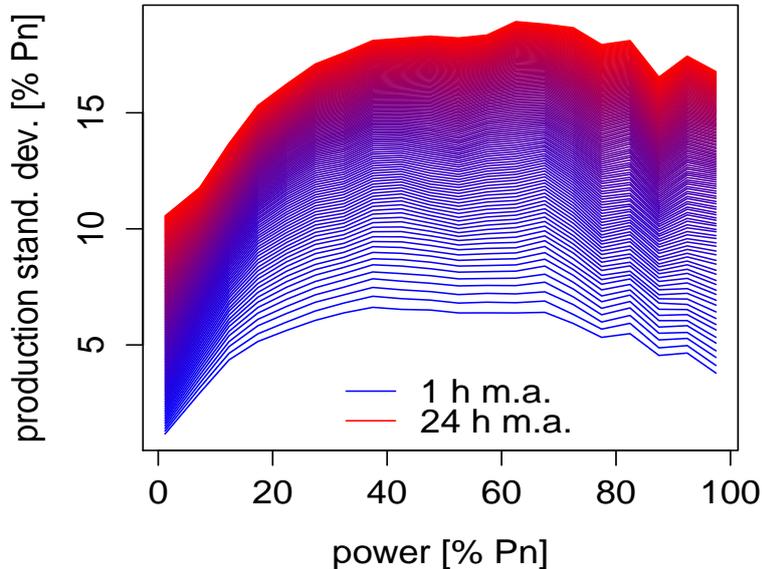
5

Figure 3: **Production standard deviation as a function of the production level** – Empirical estimates of the production standard deviation computed from sliding windows with width ranging from 1 h to 24 h, with 20 minutes' increment. One can observe how the production distribution tends to spread out, while the production level increases. The production standard deviation tends to stagnate or even decrease here, when the production level comes closer to the wind farm's nominal capacity.

The maximum duration of a null production episode is fixed at $\theta_M^1 = 20$ d. This choice is consistent with results from other work. In [22] for instance, the return period of a low wind episode of 20 days has been estimated at 1 every 10 years (in coastal areas). Then, it seems highly unlikely for an episode of null production to exceed a duration of 20 days. We assume that the maximum duration θ_M^3 of an episode with high production value is related to θ_M^1 in the same way that the associated average durations are, i.e., $\theta_M^1 = c\,\theta_M^3$. We further assume that an increasing ramp cannot be followed by a decreasing ramp in a period less than $\theta_m^3 = 1$ h. We also discard situations identified with extreme cases of *cut-off*[2] by setting $\theta_m^1 = 1$ h. Here again, values are based on our sole judgment and should be checked a fortiori.

Finally, we consider the mean duration of a ramp to have the following values $\lambda_{T_2} = 1/6, 1/2, 1, 2, 3$ h, while assuming the duration of a ramp does not exceed $\theta_M^2 = 6$ h. A ramp's minimum duration θ_m^2 is fixed at 10 min, which also corresponds to the signals' resolution chosen to simulate production. We also considered two categories of noise intensity. The first one (low noise intensity) is characterized by parameters with the value $a_1 = 2.5\,\%P_n$, $a_2 = 10\,\%P_n$ and $p_1 = 25\,\%P_n$ in Equation (5). The second one (high noise intensity) with $a_1 = 5\,\%P_n$, $a_2 = 20\,\%P_n$ and $p_1 = 25\,\%P_n$. An example of a production episode lasting 7 days simulated from our ramp model is shown in Figure 4. In the following, we will consider production episodes derived from the concatenation of $N = 100$ noisy elementary ramp profiles R^k.

---

[2]A "cut-off" refers to a situation with strong winds which result in wind turbines halting temporarily for safety purposes, causing a sudden production decrease. When the wind weakens slightly, the turbines can then restart and production quickly recovers a high level.
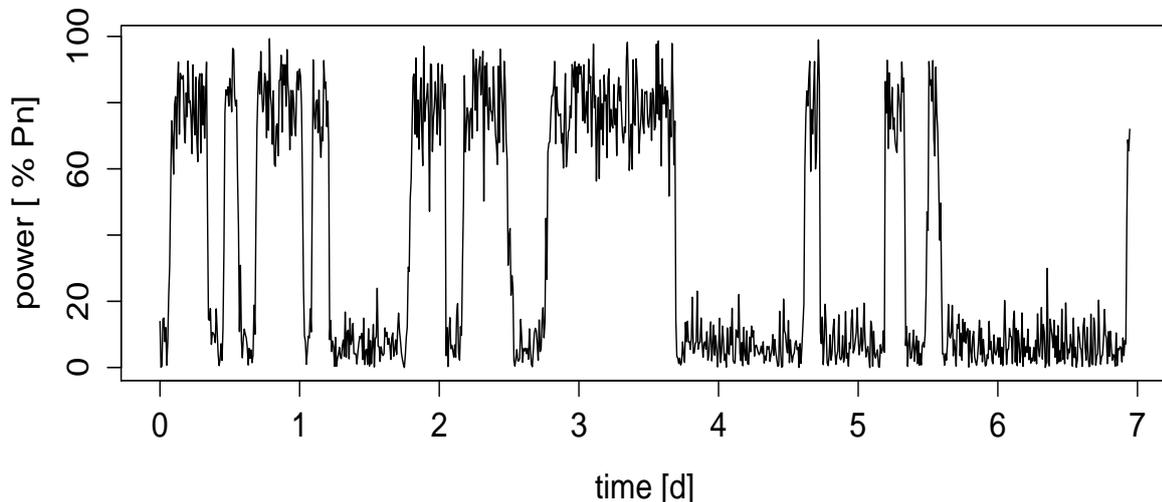
Figure 4: **Simulation example** – production episode lasting 7 days simulated from the proposed ramp model. Here, the amplitude of ramps is set at $A = 80\%P_n$. Ramps are, on average, of short duration, i.e., $\lambda_{T_2} = 1/2$ h, and rather close to one another $\lambda_{T_1} = 12$ h. The characteristic ratio between the duration of low and high production episodes has been set at $c = 2$, while noise of high intensity has been added to the ramp profile (see text for more details).

2.3 The evaluation of a ramp detection approach seen as the evaluation of a classification problem

When simulating variations of a signal, one knows exactly which variation is a ramp or not. We can count the ramps detected by a specific approach and classify the results, considering them in a supervised classification context. The particularities of ramp detection seen as a classification exercise (e.g., number and definition of classes), depends on the considered model of ramps. Criteria to consider in the evaluation of a ramp detection approach may then evolve with the ramp's model as well. The ramp model proposed in this paper only distinguishes variations due to noise from ramps in general, through associating identical properties (i.e., amplitude and duration) to both increasing and decreasing ramps. The ramp detection exercise then relates to a 2-classes classification problem.

A ramp detection approach requires measuring variations of a signal, which quantifies the sharpness and amplitude of variations. In general, it characterizes the variations of a signal over time. It is then possible to sum up this characterization to a restricted set of parameters: a particular time stamp t_I (e.g., the instant with maximum response of a derivative filter [4]), a combined measure I of the sharpness and amplitude (resp. the intensity of a variation [4]), a set of intensity measurements at different scales, or an amplitude and a blur parameter of a model estimated from the latter [11, 12]. We consider the characterization of a variation to be restricted to only two parameters: the timing t_I and intensity I of a variation. By looking at different detection approaches, we will consider different definitions of these two parameters.

If appropriate, the intensity measure used on a signal simulated from the proposed ramp model results in a two-component mixture distribution of intensity measurements. A first component is associated with low (absolute) intensity values which characterize variations due to noise. A second one is associated with high values of $|I|$ which characterize ramps. In a second step to detect ramps, a detection approach involves a thresholding procedure. By looking at a threshold intensity value $\tau > 0$, we can classify a

7

signal's variations $f_\tau : \mathcal{I} \to \{0,1\}$:

$$f_\tau(I) = \mathbb{1}_{\{|I| \geq \tau\}}.$$

Besides the detection results of an approach, it is possible to associate each variation of a simulated signal with a variable $Y \in \{0,1\}$, which denotes if such a variation is actually a ramp or not: $Y = 1$ if it is a ramp, $Y = 0$ if not. Thus, evaluating a detection approach comes down to evaluating a classification problem.

Detecting a ramp also relies on an approach's ability to accurately estimate a ramp's timing. To consider a ramp detected, a parameter δ representing the maximum admitted error in estimating such a timing, must be introduced. Our experiment involves two consecutive ramps with the same direction being separated by at least 2 h and 20 min. To avoid any confusion between two distinct ramps, we fixed the value of δ at 1 h and 10 min.

Defining the intensity of a ramp from a local maximum in a derivative signal (like it's done in [4]), is not always possible. For instance, in the case of filtering performed at a (large enough) scale, the presence of two ramps with opposite directions in the filter's support may result in a derivative with low absolute value and without any local maximum. The filter's width, which is too large relative to the time duration separating those ramps allows neither an accurate measurement nor the ability to define a characteristic intensity. However, it is still possible to define the associated classification result since the ramps in question may be considered as not detected, i.e. $f_\tau(I) = 0, \forall \tau > 0$. By convention, we may associate the intensity value $I = 0$.

We have introduced ramps' detection in a signal as a classification exercise; we can now evaluate it using criteria derived from statistical classification literature. This evaluation basically comes down to assessing the "distance" separating the two distribution components, a detection approach associates to ramps and variations due to noise intensity values. Canny's first criterion [8] may be seen as such a distance. It is defined as the expected value of a ramp intensity I_1, divided by the standard deviation of the intensity value of variations due to noise I_0:

$$SNR = \frac{\mathbb{E}[I_1]}{\text{Var}[I_0]^{1/2}} \tag{6}$$

Other distances which may be more appropriate can be found in the literature [23]:

$$\mathcal{S} = \frac{\mathbb{E}[I_1] - \mathbb{E}[I_0]}{(\text{Var}[I_1] + \text{Var}[I_0])^{1/2}} \tag{7}$$

Criteria defined in Equations (6) and (7) only require that the underlying measure of variations realized through the detection exercise are evaluated. When considering the further thresholding operation, it is then possible to define other classification-based detection criteria. Let us denote $\mathcal{E}_\tau$ the average classification (i.e., detection) error. From the Bayes formula, we get:

$$\mathcal{E}_\tau = \mathbb{P}(f_\tau(I) \neq Y) = \pi_0 q_\tau + \pi_1(1 - p_\tau), \tag{8}$$

where $\pi_i = \mathbb{P}(Y = i)$, $\pi_0 + \pi_1 = 1$ denotes the prior probabilities associated with each class of variations (noise or ramps), $q_\tau = \mathbb{P}[f_\tau(I) = 1 | Y = 0]$ and $1 - p_\tau = \mathbb{P}[f_\tau(I) = 0 | Y = 1]$. The threshold τ controls a trade-off between these two errors, since q_τ decreases, while $1 - p_\tau$ generally increases with τ. Such a trade-off can be represented using a curve parameterized by τ showing the probability p_τ depending on the probability q_τ (namely a ROC - *Receiver Operating Characteristics* - curve [24]):

$$\text{ROC} = \{(q_\tau, p_\tau), \tau > 0\}.$$

Comparing two detection approaches f_τ^1 and f_τ^2 cannot be done objectively from a given threshold value τ. Indeed, such a value would correspond to different characteristics in the respective trade-off between the two misclassification error types they define. We may however compare the respective errors obtained by each approach, with a threshold value defining an optimal classification rule $\tau_{\text{opt}} = \text{argmin}_\tau \mathcal{E}_\tau$. To that end, one may introduce a cost C, $0 < C < 1$, setting the relative importance of each

misclassification error type (replacing π_0 and π_1, respectively by $C\pi_0$ and $(1-C)\pi_1$ in Equation (8)). An alternative may be to consider the *area* under the ROC curve (denoted AUC) as a performance metric:

$$\text{AUC} = \int_0^{\sup \mathcal{I}} p_\tau q'_\tau d\tau.$$

The latter has the advantage of being objective, since with no threshold value selected a priori, it does not presume any preferable characteristic in the aforementioned trade-off. The calculus actually comes down to integrating the average classification error $\mathcal{E}_\tau$ over a set of threshold values, each is optimal with regard to different cost values C [25]. Further on in the paper, we will use the AUC criterion as it eases the comparison of different detection approaches by summing up their performances to one unique number. Of course, the consideration of other, more usual criteria (such as "false alarm" and "ramp detection" rates [26]), could also be envisaged.

3 Parametric study of different filters for the characterization of ramps

Measuring the variations of a signal necessitates repeating a differentiation operation through filtering. Derivative filters in signal-processing literature generally combine a smoothing operation with the differentiation one, so as to attenuate the noise [10]. When it comes to detecting edges, an appropriate derivative filter's shape should be related to the geometry of edges one seeks to detect [8]. Using such a filter at a particular scale provides a detection approach which can be evaluated following the framework presented in the previous section. In this section, we study the performances of different edge/ramp detection filters found in both wind-energy and signal-processing literature, following the proposed evaluation framework.

3.1 Choice of filters

We considered three filters from existing literature. The first one comes from wind-energy literature [3, 26]. It measures variations through computing the difference between the maximum and minimum level of a signal in a sliding window. Henceforth referred to as "MaxMin". The two other filters are derived from signal-processing literature. These are the Prewitt filter (referred to as "DOB"), and the first derivative of a Gaussian filter (referred to as "FDG"). Each computes finite differences combined with moving average smoothing, through a rectangular kernel for the former, and a Gaussian kernel for the latter (for a mathematical formulation of the different filters, we refer to [18]). Of the three filters considered, the only filter which has infinite support is the FDG. Nevertheless, the numerical discretization involves assimilating the FDG filter to a filter with finite support. When it comes to smoothing a signal using a Gaussian filter, truncating the latter to a width equal to four times the value of the scale parameter is often considered a reasonable choice [27]. The "wider spread" of its derivative requires truncating the FDG filter to a higher width. We chose to truncate it to a width equal to 6 times the value of the scale parameter. To ensure correct normalization whatever its width, we renormalized the FDG filter after it was truncated and discretized. In the following, we study the detection performances of the FDG, DOB and MaxMin filters for filter widths ranging from 20 min to 12 h with 20 min increment.

3.2 Different filters and their detection performance

The aim of estimating the signal-to-noise ratio (SNR, see Equation (6)) at a filter's output, is to assess a filter's ability to discriminate ramps from noise in a signal. This ratio's value increases with the difference between the filter's response to both a ramp and noise. It evolves with the filter's width, usually outlining a trade-off between noise reduction and limitation of interactions in the detection of nearby ramps. This trade-off can be observed with the DOB and FDG filters (Figure 5 shows the SNR evolution with filter's width. Results showed in this figure are average results considering a combination of all ramp's model

parameters values[3]). At first, increasing the filter's width allows to increase its response to a ramp while reducing the noise, thus resulting in higher SNR values. Then, the presence of nearby ramps in the filter's support causes a decrease in the filter's response to a ramp, resulting in a stagnation or even a decrease in the SNR. The DOB filter's detection performance is generally better than that of the FDG filter. This is due to the better adequation between the DOB filter's shape and the geometry of ramps as defined in our model. However, the FDG filter's shape makes it less sensitive to disturbances introduced by the presence of neighboring ramps. Thus, both the filters' detection performance balances at a filter's higher width values.
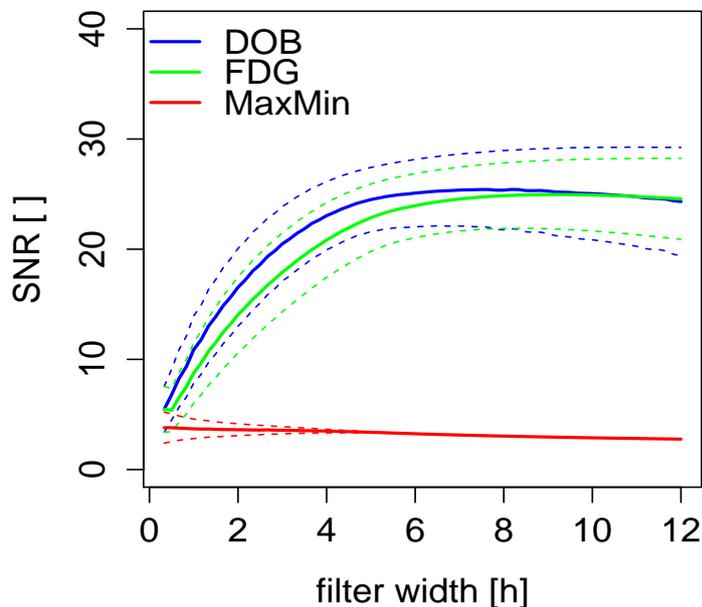

Figure 5: **Signal-to-noise ratio at filter's output (Canny's first criterion)** − Average ratio (solid lines), + / - standard deviation (dashed lines) for DOB, MaxMin and FDG filters, depending on the filter's width. Filter's width here ranges from 20 min to 12 h. Noise attenuation from the smoothing operation performed by the DOB and FDG filters allows for a rapid increase in detection performance when the filter's width increases. The presence of nearby ramps limits the response to a ramp, and ultimately results in SNR stagnation for a filter's higher width value. Due to the particular nature of its definition, the MaxMin filter's response to noise increases with the filter's width. This increase counterbalances the one associated with the filter's response to a ramp, thus resulting in an overall low level of SNR.

Due to the special nature of its definition, characterized by the absence of a smoothing operation, the MaxMin filter does not behave as usual. Its detection performance in comparison with the DOB and FDG filters is much lower (see Figure 5 red curve). When its width increases, its response to noise increases, instead of decreasing. This increase counterbalances the increase in the response to a ramp, resulting in the SNR's overall stagnation.

---

[3]In the following, we'll be discussing filters' and detection approaches' properties with regard to some of the proposed ramp's model (resp. detection approaches') parameters. Results showed are then average results computed from the combination of all other parameters values

3.3 Different filters and their localization performance

Two of the three criteria proposed by Canny are dedicated to evaluating a filter's performance in the localization of an edge. This localization is complicated by the multiplicity of maximum due to noise in the filter's response to an edge. It then becomes impossible to distinguish precisely which of the various maximum locates the edge. This multiplicity tends to disappear when the filter's width increases (see Figure 6, graphic on the left). For DOB and FDG filters, the result of the higher filter's width is more noise reduction and a lower multiplicity. In the case of the MaxMin filter, this multiplicity is decreased because of the filter's propensity to maintain its response to a maximum value[4]. Some work ackowledges an important multiplicity in the DOB filter's response (see [8] and references therein). Theoretically, it can be explained by a failure of the DOB filter to filter high frequencies. One way to look at it is to compare the *gain* of DOB and FDG filters (i.e., the modulus of their Fourier transform [28]).
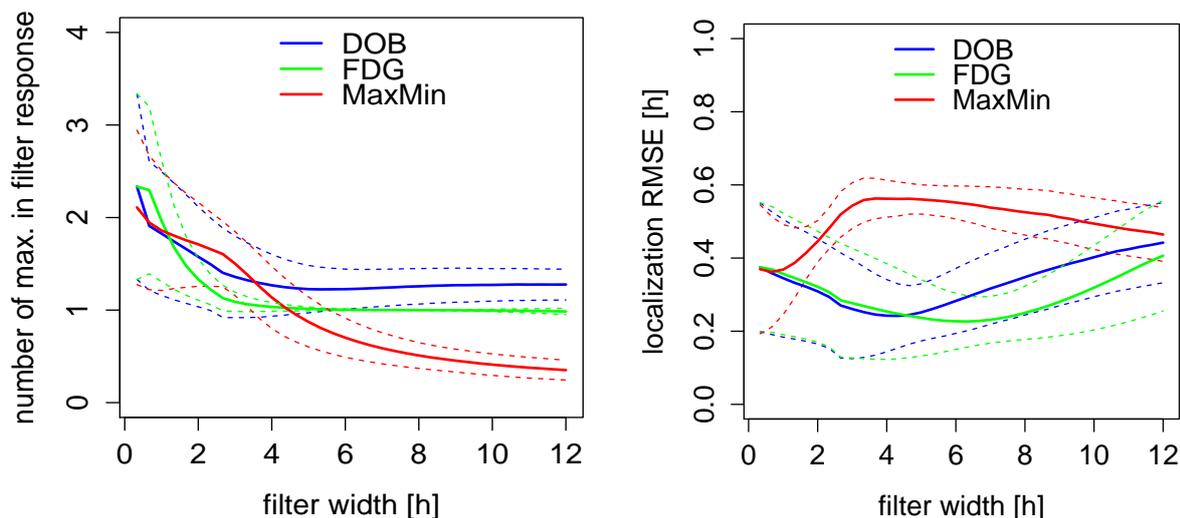

Figure 6: **Filter's responses multiplicity and ramps localization error** – Multiplicity of maximum in the response of a filter to a ramp (graphic on the left), and root mean square of the ramp localization error (graphic on the right), for the DOB, FDG and MaxMin filters, depending on the filter's width. Increasing the filter's width reduces the multiplicity of responses to a ramp. In the case of the MaxMin filter, this is not due to noise reduction, as it is for the DOB and FDG filters, but to its propensity for maintaining its maximum value. The MaxMin filter's noise sensitivity causes a rapid increase of a ramp's localization error with the filter's width. If this error decreases for filter's higher widths, this is due to the fact that errors higher than δ, have not been taken into account when computing the average localization error.

The filter's width controls, as for detection performance, a trade-off in reducing ramps' localization error. Increasing the filter's width first allows this error to be reduced through noise reduction. This phenomenon can be observed with the DOB and FDG filters (see Figure 6, graphic on the right). Then, for filter's higher widths, the benefit of reducing noise is counterbalanced by the delocalization effect produced by the presence of neighboring ramps. Because of its particular shape, i.e., less weighted at

---

[4]Typically for a large filter's width values, its response may remain constant in the interval considered here for the evaluation of the localization error. In such cases, the number of maximum associated with a response to a ramp has been set at 0, which explains the average number of maximum < 1 that can be observed in Figure 6 (graphic on the left for filter's widths > 4 h).

the extremities of its support, the FDG filter is less sensitive to it than the DOB filter. The MaxMin filter's propensity to maintain its response to a maximum value requires a suitable ramp location strategy. We chose to locate a ramp from the center of the interval for which the filter's response maintains its maximum value. The MaxMin filter does not attenuate the noise. The increase in the filter's width results in a rapid increase of the localization error of a ramp.

The use of a large width filter in the detection of nearby ramps may be responsible for a major relocation of the filter's response to each of these ramps. It sometimes results in the absence of filter response in the time interval considered around a ramp (defined by δ, see the previous Section). It would have been natural to associate a maximum localization error (i.e., equal to δ) with these situations. However, we chose to ignore them in the calculus of the root mean square ramp localization error. Thus, the influence of such situations is reflected by a reduction (instead of an increase) in the MaxMin filter localization error beyond a certain width. Nevertheless, it further serves to underline the generally poor performance of the MaxMin filter.

4   Parametric study of different multi-scale detection approaches

In this section, we study the performance of different multi-scale ramp detection approaches. The goal is to determine whether or not (and, if yes, how) the use of information collected at different scales in a variation's characterization, allows for the better detection of ramps in a signal.

4.1   State-of-the-art multi-scale detection approaches

We considered two types of multi-scale approaches. In the first case, the characterization of a variation is based on a combination of scales; while in the second case, it is based on a local scale selection procedure. In the first type of approaches, the intensity I and the timing t_I of a ramp are defined from the local maximum of a signal f_t (in absolute value), defined from the combination of a filter response at different scales:

$$f_t = K \odot_{i=1}^{n} (h_{s_i} * p)_t,$$

where $s_1 < \cdots < s_n$ are the different scales, h_{s_i} the detection filter at scale s_i, p the signal of interest, $\odot$ the algebraic operation of combination and K an appropriate normalization constant with respect to $\odot$. In [5], $\odot$ is the sum, and $K = 1/\max|\sum_i (h_{s_i} * p)_t|$. The measure of a variation f_t at a given time t is then expressed relatively to some maximum value measured in a signal. In the following, we prefer using $K = 1/n$; the measure of a variation is then expressed in absolute terms. Other measures of variation can be considered through other algebraic operations. In [15] for instance, $\odot$ indicates a product. Taking $K = 1$, and replacing f_t by f_t^{norm} by setting $f_t^{\mathrm{norm}} = \mathrm{sign}[(h_{s_n} * p)_t].|f_t|^{1/n}$, then allows the units of measurement to be kept. In the following, we denote the two considered approaches (sum and product) by the names of the authors who proposed them, i.e. "Gallego" and "Bao" respectively.

In the second type of approaches, we consider an approach proposed in [29], henceforth "Lindeberg". In this approach, a ramp's intensity and timing are defined from the local maximum of the filtered signal (in absolute value), at a scale chosen according to the following:

$$s_{\mathrm{loc}} = \operatorname*{argmax}_{0 < s \leq s_{\max}} |I(s)|.$$

The latter is defined only if a local maximum associated with a variation in the derivative signal (in absolute value) exists for all scale values below a maximum scale $s_{\max} \leq s_n$. The FDG filter ensures such an existence, unlike the DOB and MaxMin filters (see [18] for a counter example with real production data). More generally, such an existence is related to a property sometimes referred to as *causality* in the literature of a signal's time-scale representations [30, 31]. The causality property means that some elements describing the structure of a signal disappear as the scale gets coarser, leaving only those offering a more synthetic representation. According to some work, the Gaussian kernel is probably the only one from a broad class of operators to offer this property [31, 32].

In what follows, we study the performance of the different approaches based on the use of a continuous wavelet transform with the first derivative of a Gaussian as a mother wavelet. The normalization assumed through the use of a wavelet transform (L2 normalization) must have an influence on the performance of certain approaches (e.g., through the scale selection procedure defined in the Lindeberg approach). Nevertheless, such an influence is not studied here and should be further investigated.

The use of multiple scales does not depart from the choice of a minimum and a maximum scale. The smallest scale s_1 is defined here by the resolution of the signals in accordance with the choice made in [5]. The scale increment is chosen three times lower than that of the resolution of the signals. Finally, the maximum scale s_n is a parameter whose influence on the different approaches' performance will be investigated. We consider the following values: $s_n = 1/3, 1/2, 1, 1.5, 2, 2.5, 3$ h.

4.2 Detection performance

The influence of ramps' closeness: Results from the estimation of the AUC, depending on the average duration of an episode of low production λ_{T_1} and the ratio defining the average duration of an episode of high production c, are shown in Figure 7 (graphs on top). These results were obtained from a maximum scale value $s_n^\star$ guaranteeing the best possible detection performance, i.e. $s_n^\star(\lambda_{T_1}, c) = \text{argmax}_{s_n} \text{AUC}(\lambda_{T_1}, c)$. They highlight a difference between the detection performances of approaches based on scales' combination (i.e., Gallego and Bao), and the approach involving a local scale selection procedure (i.e., Lindeberg). Indeed, the performance of the latter is maintained at a high level, regardless of the characteristics defining the time duration separating two consecutive ramps in a signal. On the other hand, the first two approaches reveal a declining performance when the average time duration that separates ramps, decreases.

AUC's usually high value makes it difficult to quantitatively appreciate the performance gap between different approaches. This is a inconvenience sometimes acknowledged in the statistical literature [25]. Here, it must be related to the simplicity of the considered detection exercise. Indeed, our model does not allow to reproduce properly the numerous textures that appear in real wind power production signals and that may complicate the detection of ramps. To better appraise the difference in the various approaches' performances, we made additional simulations considering shorter average duration of an episode of low production λ_{T_1}, i.e., $\lambda_{T_1} = 1, 2, 3, 4, 5$ h. We realize that such values may be unrealistic. Through such a procedure however, we may expect to get closer to the difficulty there is to detect ramps on real wind power production signals. The results can be seen in Figure 7 (graphs at the bottom, please note the changing scale in the ordinate axis). They confirm our initial observations.

The comparison of these performances with those obtained from the use of a single globally optimal scale, i.e., $s_{\text{opt}}(\lambda_{T_1}, c) = \text{argmax}_s \text{AUC}(\lambda_{T_1}, c)$, can be well appreciated here. This comparison clearly discriminates between the approaches based on scales' combination or on a local scale selection procedure. On one hand, the use of a single scale s_{opt} in the detection of all ramps of a signal provides better results than a strategy based on the combination of different scales, as in the Bao or Gallego approach. On the other hand, a local procedure of scale selection, as in the Lindeberg approach, allows for detection performance to increase (or at least to maintain) as regards the use of a single globally optimal scale.

The performance gain of a multi-scale approach with respect to another approach is only effective if the scale $s_n^\star$ can be determined a priori. This requires knowing the value of λ_{T_1} and c parameters. A rough estimate of these parameters made from real signals' analysis can lead to a poor choice in using the value of the parameter s_n. It is thus desirable to have detection performances that are not too sensitive to the value of s_n. Figure 8 shows the decline in detection performance for the Gallego (left graphs) and Lindeberg (right graphs) approaches, with respect to their respective optimal performance (i.e., $\text{AUC}(s_n^\star)$), depending on the value of s_n. This decline is represented here as a function of the average inter-ramps' duration (the different curves correspond to the different values of the parameter λ_{T_1}, graphs on top are for values $\lambda_{T_1} = 6, 12, 24, 48, 96$ h while those at the bottom are for values $\lambda_{T_1} = 1, 2, 3, 4, 5$ h, c has been set at $c = 5$). When the signal ramps are sufficiently far from one another, the value of s_n has little effect on both the performances of the Gallego and the Lindeberg approaches. However, if consecutive ramps are close (i.e., $\lambda_{T_1} \leq 3$ h, graphs at the bottom, again please note the changing scale in the ordinate axis), the choice of s_n becomes critical to ensure the Gallego approach to maintain detection
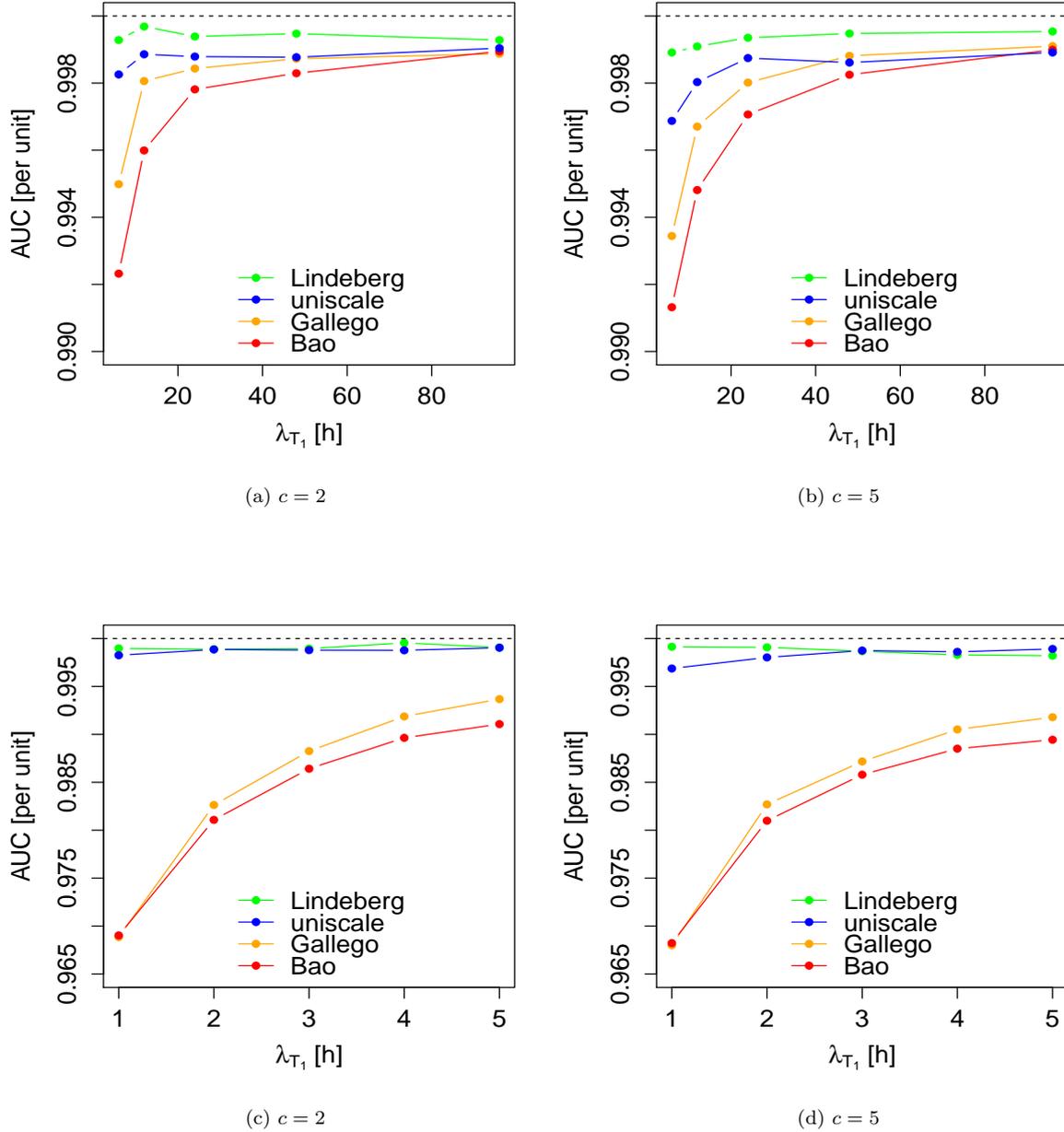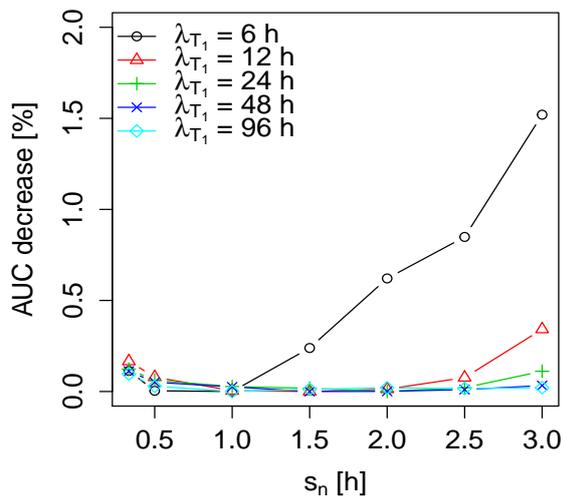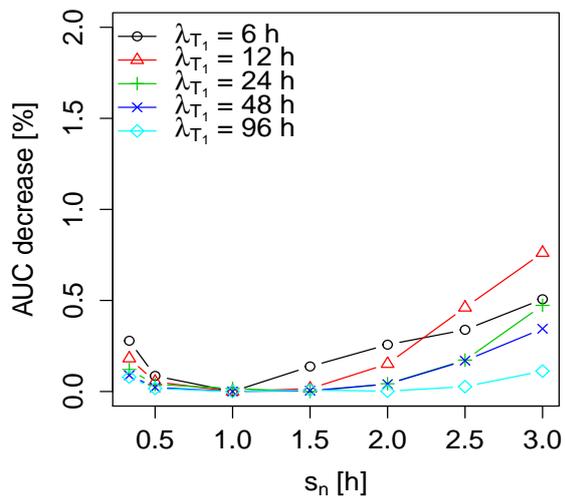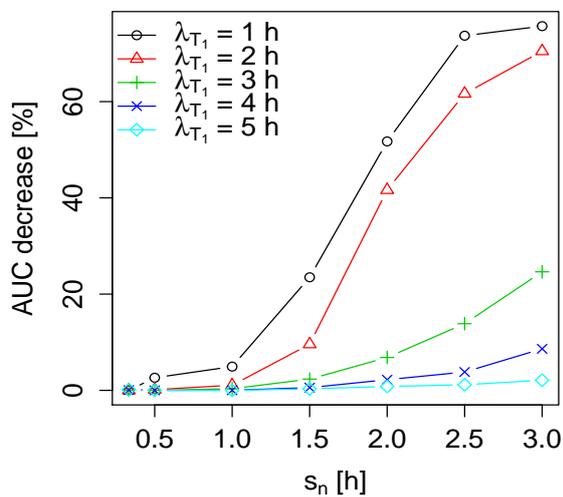
Figure 7: **Optimal detection performance depending on average inter-ramps' duration** – Maximum AUC (i.e., for a maximum scale parameter value $s_n^\star = \mathrm{argmax}_{s_n} \mathrm{AUC}(\lambda_{T_1}, c)$), depending on the average duration of an episode of low production λ_{T_1} and the ratio defining the average duration of an episode of high production c, for multi-scale approaches: Gallego, Bao and Lindeberg. Graphs on top are for parameter values $\lambda_{T_1} = 6, 12, 24, 48, 96$ h, while graphs at the bottom correspond to values $\lambda_{T_1} = 1, 2, 3, 4, 5$ h. Performance when using a single globally optimal scale for detecting all the ramps in a signal (blue curve), is here given as a reference. It allows for a clear discrimination between an approach's performance based on a local scale selection procedure (i.e., Lindeberg), and those based on scales' combination (i.e., Bao and Gallego).
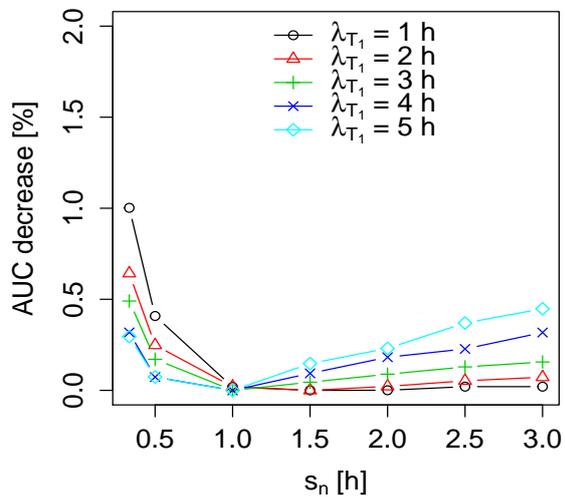
14

Figure 8: **Detection performance sensitivity to the choice of s_n** − Decrease in detection performance with respect to the optimal ones (i.e. $\mathrm{AUC}(s_n^\star)$), for the Gallego (left graphs) and Lindeberg (right graphs) approaches, depending on the value of s_n. Graphs on top are for parameter values $\lambda_{T_1} = 6, 12, 24, 48, 96$ h, while graphs at the bottom correspond to values $\lambda_{T_1} = 1, 2, 3, 4, 5$ h. When considering close consecutive ramps, i.e. $\lambda_{T_1} \leq 3$ h (graphs at the bottom), one can see how the performance of the Gallego approach is sensitive to the choice of s_n. The results presented here were obtained for a ratio $c = 5$ between the average duration of an episode of low production, and the average duration of an episode of high production.

15

performance close to the optimal ones. On the other hand, the performance of the Lindeberg approach does not seem very sensitive to the choice of s_n.

The influence of the signal-to-noise ratio - the case of nearby ramps ($\lambda_{T_1} = 6$ h and c = 5):
We now focus on challenging detection situations, namely on signals in which consecutive ramps are close to one another. One can have more insight into the different approaches' performance by looking at the characteristics of the distributions of the intensity values, associated with each class of variations (ramps or noise). Considering sample estimates of the average intensity values $\overline{I}_0$ and $\overline{I}_1$ and standard deviation $\hat{\sigma}_0$ and $\hat{\sigma}_1$ associated with each class of variations, one can define an estimate $\hat{\mathcal{S}} = (\overline{I}_1 - \overline{I}_0)/(\hat{\sigma}_1^2 + \hat{\sigma}_0^2)^{1/2}$ of the criterion given in Equation (7). Values close to 0 indicate little discrimination of ramps in a signal, the higher the better.

Table 1 shows estimates obtained in the case of nearby ramps: for different values of amplitude A and average time duration λ_{T_2} of a ramp, as well as for different intensity j of noise. The combination of different ramps amplitude and duration values with different noise intensities, can be identified as different signal-to-noise ratios. A detection approach performs better when this ratio increases. Some of the detection approaches considered are very sensitive to this ratio. One can see that the Bao and Gallego approaches have good potential for discriminating ramps in a signal when this ratio is high (i.e., for high values of A associated to low values of λ_{T_2} and j), even when consecutive ramps are close to one another. However, these approaches' performance sharply decreases when the signal-to-noise ratio drops. Yet, the Lindeberg approach is less sensitive to it, and it performs better for the most difficult detection cases (see Table 1 **(c)**).

Table 1: **Characteristics of the distributions of the intensity values associated with each class of variation** – Sample estimates of the mean and standard deviation of the intensity value associated with each class of variation (ramps of noise), along with estimates of the $\mathcal{S}$ criterion (see Equation (7)). The results shown here were obtained for a maximum scale parameter value $s_n = s_n^\star$, parameters characterizing nearby ramps $\lambda_{T_1} = 6$ h and c = 5, different amplitudes A and average durations λ_{T_2} of ramps, and for different intensities of noise j. The combination of these last three parameters defines a signal-to-noise ratio (**(a)**, **(b)** and **(c)** describes decreasing ratios). If the Gallego and Bao detection approaches are potentially good for discriminating ramps when this ratio is high, their performance sharply decreases with it. However, the Lindeberg approach shows superior performances in the most difficult detection cases (e.g. **(c)**).

(a) $A = 100\% \ P_n$, $\lambda_{T_2} = 10$ mn, $j = 1$

	$\overline{I}_1$ [%$P_n.\sqrt{h}$]	$\overline{I}_0$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_1$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_0$ [%$P_n.\sqrt{h}$]	$\hat{\mathcal{S}}$ []
Gallego	62.9	1.2	5.5	0.8	11.2
Bao	51.6	1.4	1.6	0.9	**27.3**
Lindeberg	75.3	2.0	4.7	2.5	13.7

(b) $A = 70\% \ P_n$, $\lambda_{T_2} = 1$ h, $j = 1$

	$\overline{I}_1$ [%$P_n.\sqrt{h}$]	$\overline{I}_0$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_1$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_0$ [%$P_n.\sqrt{h}$]	$\hat{\mathcal{S}}$ []
Gallego	43.0	1.7	6.7	3.1	5.6
Bao	28.7	2.3	8.3	3.0	3.0
Lindeberg	59.4	2.1	5.1	1.9	**10.6**

(c) $A = 50\% \ P_n$, $\lambda_{T_2} = 3$ h, $j = 2$

	$\overline{I}_1$ [%$P_n.\sqrt{h}$]	$\overline{I}_0$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_1$ [%$P_n.\sqrt{h}$]	$\hat{\sigma}_0$ [%$P_n.\sqrt{h}$]	$\hat{\mathcal{S}}$ []
Gallego	22.2	5.4	6.1	5.3	2.1
Bao	14.4	3.9	4.9	2.9	1.8
Lindeberg	34.7	4.1	6.6	2.8	**4.2**

4.3 Localization performance

In addition to detection performance, we studied the performance of multi-scale detection approaches in locating well ramps. The results contrast with those obtained in terms of detection. Figure 9 shows the ramp localization root mean squared error (RMSE) for different detection approaches, depending on the parameters controlling the in-between-ramps' duration, λ_{T_1} and c. The results shown here were obtained from a maximum scale parameter value s_n minimizing the localization RMSE. The Gallego approach has the best localization performance, especially for close ramps. It must be taken into account that this result cannot be generalized to any other detection approach based on scales' combination. Indeed, the Bao approach has a very poor localization performance.

The choice of an appropriate maximum scale parameter value s_n is difficult, because a suitable value for detection is not necessarily adapted in terms of localization. Our simulations showed that to keep localization performances close to optimal, s_n should stay around 1 and 1.5 h for the Lindeberg approach. These values seem particularly suitable as they also allow good detection performance (see Figure 8). In the case of the Gallego approach, choosing s_n so as to keep performance satisfactory both in terms of detection and localization may be more difficult. Indeed, if low values of s_n (e.g., $s_n = 0.5$ h or 1 h) are adapted in terms of detection, it is necessary to use larger values (e.g., between 1.5 h and 2 h) to optimize localization performance.
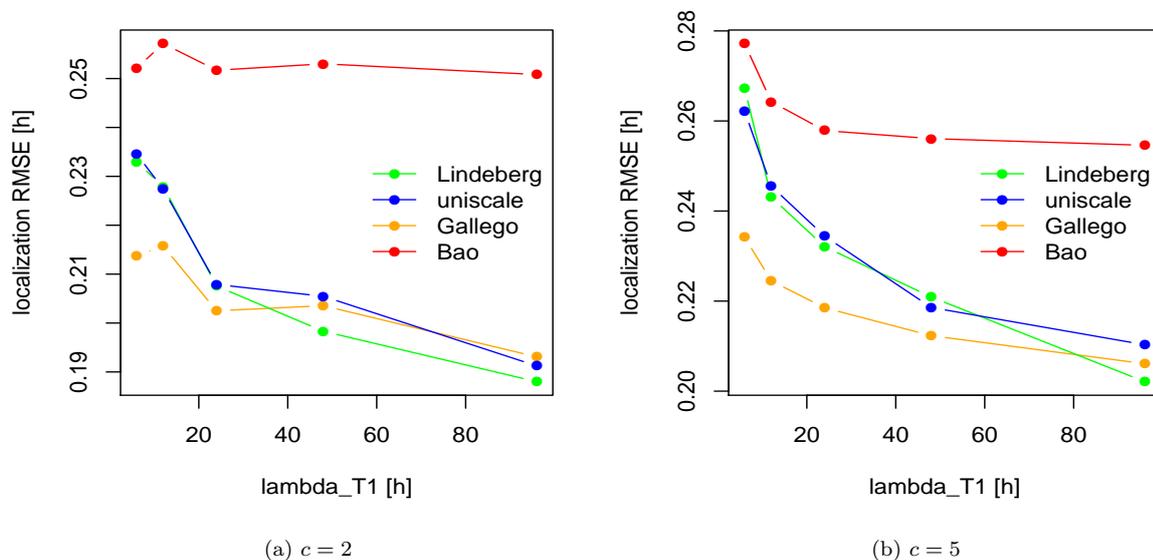

(a) $c = 2$ (b) $c = 5$

Figure 9: **Optimal localization performances depending on average inter-ramps' duration** – Ramp localization root mean squared error of the considered detection approaches, depending on the inter-ramps' duration parameters λ_{T_1} and c. These results contrast those obtained for detection, since it is the Gallego scales' combination approach which performs the best for close consecutive ramps.

5 Conclusion

In wind-energy literature, the study of large and sharp wind-power variations has seen new characterization and detection approaches emerge. However, there is no existing work dedicated to their evaluation using tangible quantitative criteria, to establish these approaches' validity. In this paper, we have proposed a framework from which these approaches can, indeed, be evaluated. We have proposed a theoretical

model to formally define the "large" and "sharp" character of wind-power ramps. We have also described quantitative criteria to objectively evaluate different approaches to detecting ramps as defined by our model.

The ramps' model we propose includes some of the essential aspects already described in edge-detection literature. We have pushed the model further so as to adequately represent certain properties of the wind-power production process. By modeling the time between two consecutive ramps using random variables, we tried to reproduce the random nature with which such ramps follow one another in real production signals. Through noise modeling, we took care to reproduce some of the non-stationary characteristics and bounded aspect of the wind-power process. Besides the usual criteria proposed by Canny [8], we described some criteria derived from statistical-classification literature which can be used to evaluate a detection approach's performance.

From the proposed ramp model and evaluation criteria, we studied different filters and multi-scale detection approaches found in wind-energy literature. We compared their performance with filters and other multi-scale detection approaches from edge-detection literature. The results of our study show that the filter proposed in [3] is very sensitive to noise and has a lower performance than more conventional filters. They also show that some of the multi-scale approaches proposed in the literature which are based on scales' combination, may have difficulties in maintaining a high performance level in detecting ramps when those are closely surrounded by other ramps or by other elements in a signal. Moreover, their performances may be very sensitive to the choice of the highest scale used for combination which then makes the interest of using those approaches less obvious. Finally, other multi-scale approaches like the ones based on local scale selection procedures for instance, may be preferable.

Of course, the proposed ramp's model upon which rely our observations is limited. Nevertheless, we believe the work proposed here is a first step towards systematically evaluating ramp characterization approaches. More insight could be obtained about them if the proposed ramp model is further developed to better integrate some characteristics of the wind power process (e.g., better distribution and auto-correlation modeling, etc.,), and is able to reproduce important textures that are observed in real production signals.

Acknowledgements

References

[1] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. The state of the art in short-term prediction of wind power - A literature overview, 2nd edition. Technical report, SAFEWIND EU project, deliverable report Dp-1.4. [Available online: http://www.safewind. eu], 2011.

[2] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6):1725 – 1744, 2008.

[3] N. J. Cutler, M. Kay, K. Jacka, and T. S. Nielsen. Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT. *Wind Energy*, 10(5):453–470, 2007.

[4] A. Bossavy, R. Girard, and G. Kariniotakis. Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy*, 16(1):51–63, 2013.

[5] C. Gallego, A. Costa, . Cuerva, L. Landberg, B. Greaves, and J. Collins. A wavelet-based approach for large wind power ramp characterisation. *Wind Energy*, 16(2):257–278, 2013.

[6] M. Petrou and J. Kittler. Optimal edge detectors for ramp edges. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(5):483 –491, may 1991.

[7] D. J. Williams and M. Shah. Edge characterization using normalized edge detector. *CVGIP: Graphical Models and Image Processing*, 55(4):311 – 318, 1993.

[8] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679 –698, 1986.

[9] J. Prewitt. Object enhancement and extraction. *Picture Process. Psychopict*, pages 75–149, 1970.

[10] D. Ziou and S. Tabbone. Edge detection techniques - An overview. *International Journal of Pattern Recognition and Image Analysis*, 8:537–559, 1998.

[11] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(7):710 –732, jul 1992.

[12] C. Ducottet, T. Fournel, and C. Barat. Scale-adaptive detection and local characterization of edges based on wavelet transform. *Signal Processing*, 84(11):2115 – 2137, 2004.

[13] D. Demigny. On optimal linear filtering for edge detection. *Image Processing, IEEE Transactions on*, 11(7):728 – 737, 2002.

[14] O. Laligant, F. Truchetet, and F. Meriaudeau. Regularization preserving localization of close edges. *Signal Processing Letters, IEEE*, 14(3):185 –188, march 2007.

[15] P. Bao, L. Zhang, and X. Wu. Canny edge detection enhancement by scale multiplication. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(9):1485 –1490, sept. 2005.

[16] J.H. Elder and S.W. Zucker. Local scale control for edge detection and blur estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(7):699 –716, jul 1998.

[17] Y.H. Wan and D.Jr. Bucaneg. Short-term power fluctuations of large wind power plants. *Journal of Solar Energy Engineering*, 124(4):427–431, 2002.

[18] A. Bossavy. *Characterization and probabilistic forecasting of wind power ramps.* PhD thesis, MINES ParisTech. [*In french.* Available online at: http://pastel.archives-ouvertes.fr/docs/00/80/32/34/PDF/2012ENMP0058.pdf], 2012.

[19] H. Bludszuweit, J.A. Dominguez-Navarro, and A. Llombart. Statistical analysis of wind power forecast error. *Power Systems, IEEE Transactions on*, 23(3):983 –991, aug. 2008.

[20] P. Pinson. Very-short-term probabilistic forecasting of wind power with generalized logit - normal distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(4):555–576, 2012.

[21] A. Lau and P. McSharry. Approaches for multi-step density forecasts with application to aggregated wind power. *The Annals of Applied Statistics*, 4(3):1311–1341, 2010.

[22] P.G. Leahy and E.J. McKeogh. Persistence of low wind speed conditions and implications for wind power variability. *Wind Energy*, pages n/a–n/a, 2012.

[23] C. Marzban. The ROC curve and the area under it as performance measures. *Weather and Forecasting*, 19:1106–1114, 2004.

[24] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861 – 874, 2006.

[25] D.J. Hand. Measuring classifier performance: acoherent alternative to the area under the ROC curve. *Machine Learning*, 77:103–123, 2009.

[26] B. Greaves, J. Collins, J. Parkes, and A. Tindal. Temporal Forecast Uncertainty for Ramp Events. *Wind Engineering*, 33(11):309–319, 2009.

[27] D. Demigny, L. Kessal, and J. Pons. Fast recursive implementation of the Gaussian filter. In *11th IFIP International Conference on Very Large Scale Integration*, number 1, pages 339–346, Montpellier, France, december 2001. LIRMM.

[28] G. Blanchet and M. Charbit. *Traitement numrique du signal*. Hermes, 1998.

[29] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–154, 1998.

[30] T. Lindeberg. Scale-space for discrete signals. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(3):234 –254, mar 1990.

[31] J. Bigot. *Recalage de signaux et analyse de variance fonctionnelle par ondelettes. Applications au domaine biomdical*. PhD thesis, Universit Joseph-Fourier-Grenoble 1. [Available online : http://tel.archives-ouvertes.fr/docs/00/04/60/08/PDF/tel-00004362.pdf], 2003.

[32] J. Babaud, A.P. Witkin, M. Baudin, and R.O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(1):26 –33, jan. 1986.