Effect of modulation depth, frequency, and intermittence on wind turbine noise annoyance[a)]

Christina Ioannidou,[b)] Sébastien Santurette, and Cheol-Ho Jeong

Department of Electrical Engineering, Technical University of Denmark, DTU Bygning 352, Ørsteds Plads, 2800 Kgs. Lyngby, Denmark

Amplitude modulation (AM) may be an important factor for the perceived annoyance of wind turbine noise (WTN). Two AM types, typically referred to as "normal AM" (NAM) and "other AM" (OAM), characterize WTN AM, OAM corresponding to having intermittent periods with larger AM depth in lower frequency regions than NAM. The extent to which AM depth, frequency, and type affect WTN annoyance remains uncertain. Moreover, the temporal variations of WTN AM have often not been considered. Here, realistic stimuli accounting for such temporal variations were synthesized such that AM depth, frequency, and type, while determined from real on-site recordings, could be varied systematically. Listening tests with both original and synthesized stimuli showed that a reduction in mean AM depth across the spectrum led to a significant decrease in annoyance. When the spectrotemporal characteristics of the original far-field stimuli and the temporal AM variations were taken into account, the effect of AM frequency remained limited and the presence of intermittent OAM periods did not affect annoyance. These findings suggest that, at a given overall level, the AM depth of NAM periods is the most crucial AM parameter for WTN annoyance. © 2016 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4944570]

[JFL] Pages: 1241–1251

I. INTRODUCTION

One inherent characteristic of wind turbine noise (WTN) is amplitude modulation (AM), i.e., temporal variations in noise level occurring at an approximately constant period (quasi-periodic noise).[1] As this variation in level repeats every time a rotating blade passes a certain point of the blade trajectory, the perceived noise is amplitude-modulated at the blade passing frequency, typically between 0.5 and 2 Hz for most modern large scale wind turbines.[2,3]

WTN has often been considered to be a relatively annoying source and resulting complaints are typically related to audible broadband aerodynamic sound.[1] Although AM of WTN was at first believed to be only perceivable at regions close to a wind turbine, it has been shown to be perceivable at longer distances if the background noise level is quite low.[1,4]

Initially, it was found that WTN AM, sometimes referred to as "swish" or "thump," mainly occurred in spectral regions between 500 Hz and 2 kHz.[5,6] van den Berg distinguished between these two terms depending on the frequency region where the most prominent modulations occur: (1) The swishing sound, heard close to a turbine and related to the sound modulation from middle to high frequencies; and (2) the thumping sound, referring to an impulsive sound with a rapid rise time.[6,7] Two AM types were later defined based on the strength of modulations and the frequency region in which these occur (see Table I) and consistently with van den Berg's discrimination of noise types:[8]

(1) *Normal AM (NAM)*, commonly described as (blade) swish, is an inherent feature of WTN. Its maximum modulation depth occurs between 400 and 1000 Hz and it varies from 3 dB (close to the source) to about 6 dB (more than 1 km from the source). It is persistent for long periods. As the distance from the wind turbine increases, the audibility of NAM decreases.[3]

(2) *Other AM (OAM)*, subjectively described as thump, can be expressed as a change of the characteristics and spatial distribution of NAM. It is not a common, but rather intermittent feature of wind farms and its causal mechanisms are unclear. For a sound that includes OAM periods, AM occurs at lower frequencies than NAM, leading to a maximum modulation depth between 300 and 400 Hz. Furthermore, the modulation depth can reach values between 6 and 12 dB and it can be observed even at large distances (more than 1 km from the turbine).[2,3,8]

Zwicker first proposed that annoyance calculations should take into account not only loudness but also AM depth.[9] Then, van den Berg found that people living close to a wind farm (at distances between 500 and 1000 m from it) were annoyed mainly during the night because of a low-pitch thumping sound, periodic at the blade passing frequency.[6] The fact that amplitude-modulated sounds cause greater annoyance than unmodulated sounds was already observed earlier in other contexts. One study concerning the subjective effect of AM in diesel-engine exhaust noise showed that the presence of AM in this type of noise played

---

TABLE I. Spectral regions in which the most prominent AM occurs and modulation depths of NAM and OAM for observations at large distances (more than 1 km), according to Ref. 2.

Characteristics	NAM	OAM
Main spectral region	400–1000 Hz	300–400 Hz
AM depth at large distances	Up to 6 dB	6–12 dB

an important role in determining annoyance.[10] Moreover, another study focusing on the annoyance rating of low-frequency sounds showed that amplitude-modulated broadband noise was more annoying than unmodulated noise.[11]

In the specific context of WTN, Lee *et al.* showed that annoyance increased most significantly with increasing equivalent A-weighted sound pressure level (L_{Aeq}) and to a lesser extent with increasing modulation depth.[12] The same group found that amplitude-modulated WTN recorded at the near-field was more annoying than steady WTN.[13]

Another recent study focused on the subjective impact of normal and other AM on annoyance by using an appropriately designed WTN model to generate the stimuli.[14] The results agreed with previous literature,[12] showing that annoyance depends predominantly on L_{Aeq} and to a lesser extent on modulation depth and frequency.[14] Specifically, it was shown that the response to the noise was not significantly affected by the frequency content of the modulated noise (NAM or OAM), once the A-weighted sound pressure level was taken into account. Moreover, it was found that the temporal (shape and width) and spectral (frequency skew and bandwidth) parameters of the modulation pulse did not have a significant effect on annoyance.

One limitation of the latter study was that the effect of WTN on annoyance was investigated with steady AM sounds (i.e., constant AM depth and amplitude over time), while the modulation depth and spectral characteristics do vary significantly during original WTN recordings, on time scales as short as a few seconds. Moreover, AM of WTN is an intermittent phenomenon. Such factors were, so far, not taken into account but may have a significant impact on annoyance.

Therefore, the present study investigated the effect of WTN AM on annoyance via listening experiments in which the temporal variations and intermittence of the AM phenomena were taken into account. Both original and synthesized WTN stimuli were used in the experiments. The original stimuli were first processed to obtain the AM metrics of interest (e.g., AM depth and frequency), which were used to develop a model of amplitude-modulated WTN. Sets of synthesized stimuli with varying AM parameters generated via this model were used in the listening experiments to investigate the effect of AM depth, frequency, and type (NAM or OAM) on annoyance when AM temporal variations and intermittence are taken into account. Three listening experiments were designed to systematically investigate how annoyance was affected by AM depth (experiment 1), AM frequency (experiment 2), and the interaction between AM type and AM depth (experiment 3).

II. EXPERIMENTAL SETUP AND PROCEDURE

The listening experiments were performed in a double-walled, sound-attenuating, and electromagnetically shielded booth. A total of 19 participants (9 males, 10 females) aged between 23 and 28 yr were recruited by the Technical University of Denmark to participate in the experiments. All participants had pure-tone hearing thresholds below 20 dB hearing level at all audiometric frequencies and no reports of hearing difficulties and could thus be considered as having normal hearing. A subset of the 19 participants in the study carried out each experiment, such that a total of 10, 14, and 13 listeners participated in experiments 1, 2, and 3, respectively. The stimuli were presented via Sennheiser HDA200 headphones connected to a PC with MATLAB software and an RME DIGI 96/8 24-bit soundcard. Such a system reproduced only the audible signal and not the vibrations. A sampling frequency of 44.1 kHz was used for stimulus playback. The approximate A-weighted sound level was 62 dBA for all stimuli. All experiments were approved by the Science Ethics Committee of the Capital Region of Denmark (reference H-3-2014-004).

Listening experiments using stimuli with equivalent sound levels varying from 30 to 45 dBA L_{Aeq} have already demonstrated that the equivalent sound pressure level significantly contributes to noise annoyance.[2,12,13] Therefore, a fixed overall level of 60 dB sound pressure level (SPL) was used in all listening experiments. The stimulus duration in the listening experiments of Lee *et al.* and von Huenerbein *et al.* was 30 s.[2,13] The selection of this duration was based on studies concluding that a stimulus length as short as 30 s can give comparable results to long stimuli.[15,16] Therefore, the standard duration of all stimuli for this study was selected to be equal to 30 s.

The instructions given to all listeners before each experiment stated that they should imagine that they were sitting in their garden after a busy day, trying to relax. While doing so, they were asked to listen to original or synthesized WTN samples and to rate how annoying these sounds would be if they were heard continuously at their garden. They were not given any definition of annoyance. The listeners could not interrupt the sound reproduction, such that they had to listen to the whole 30-s stimulus before rating it. However, they were allowed to repeat the stimulus presentation as many times as they wished if they felt it necessary to judge its annoyance. It should be stressed that such an experimental setting was artificial and that, despite the instructions given to the listeners, ecological validity would require comparison with a field study. Here, the results should be seen as reflecting the effect of AM parameters in a laboratory environment.

After each stimulus presentation, the listeners were asked to rate the annoyance of the sample they just heard by adjusting a slider shown on a computer screen to any real value between 0 and 10. The rating scale ranged from 0 ("not annoying at all") to 10 ("unbearably annoying"). The rating terms written on the screen were "not annoying at all," "slightly annoying," "moderately annoying," "very annoying," and "unbearably annoying," corresponding to the

slider positions indicating annoyance of value 0, 3, 5, 7, and 10. The listeners could also see the rating value they gave to each sample in a small box. To derive an average estimate of annoyance, each sound sample was rated twice by each listener in experiment 1 and three times in experiments 2 and 3. For each repetition, the stimuli were presented in random sequence to avoid possible order bias. Before each experiment, a training session with 16, 8, and 4 samples for experiments 1, 2, and 3, respectively, was carried out to make the listeners familiar with the sounds they would be listening to. The total duration of experiments 1, 2, and 3, including short breaks and training, was 75, 60, and 20 min, respectively.

III. EXPERIMENT 1: EFFECT OF AM DEPTH ON ANNOYANCE

This experiment, based on a concept by Lee *et al.*,[13] investigated the annoyance caused by the combination of far-field amplitude-modulated WTN recordings and artificial unmodulated WTN. The idea was to add various levels of unmodulated WTN to the original amplitude-modulated recordings in order to systematically reduce their modulation depth. The purpose was to investigate whether AM depth has a significant effect on WTN annoyance.

A. Method

Eight amplitude-modulated WTN samples including periods of interest (i.e., varying AM depth intervals, AM intermittence intervals) were extracted from four original WTN recordings. All original samples were low-pass-filtered with a 6.5-kHz cutoff frequency. Such filtering led to samples that sounded realistic by excluding any recording noise that could affect annoyance.

For each sample, artificial unmodulated WTN was then created by filtering white Gaussian noise such that it had similar spectral characteristics to the spectrum of the original sample. The magnitude of the original sample (Fig. 1, grey color) was obtained by applying the Fourier transform and then a polynomial was fitted to its spectrum to design a filter with a magnitude response corresponding to the spectrum of the original sample. Then, white Gaussian noise was generated and filtered accordingly in the spectral domain. The filtering was done by multiplying the derived polynomial curve with the Fourier transform of the white Gaussian noise. Subsequently, the final time-domain signal was obtained via the inverse Fourier transform. Finally, the unmodulated noise was low-pass-filtered below 6.5 kHz as the original sample and then added to the original sample. Different amounts of this unmodulated noise were added to the original sample, thus reducing its AM depth to different degrees. Then it was used to design a filter having the same magnitude response.

1. Modulation depth spectrum (MDS)

The variations in SPL over time can be used to quantify the strength of modulation by calculating the overall modulation depth (MD), i.e., the mean value of the differences of the peak-to-trough SPL values for a time step equal to the
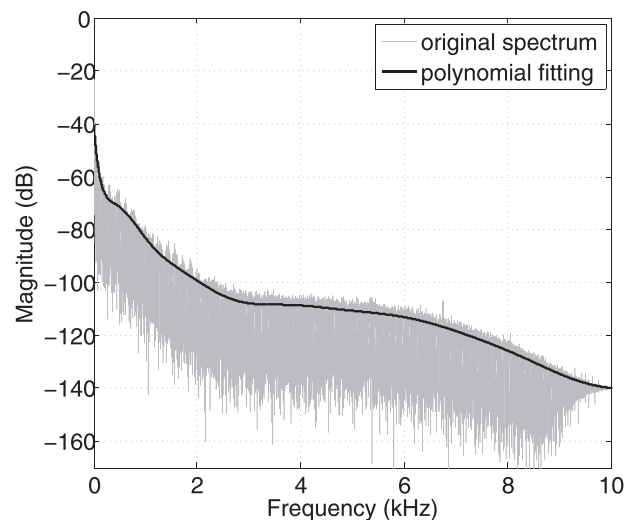

FIG. 1. Spectrum of original WTN sample 1 (grey line). The spectrum is approximated by fitting a polynomial to that spectrum (black line). That curve was used to design a filter.

blade passing period. However, Lee *et al.* pointed out that this metric is not optimal to quantify AM strength when it is only calculated for the frequency range where the SPL presents its highest values.[13] Therefore, they suggested using the *modulation depth spectrum* (*MDS*), defined as the difference between the maximum and the minimum values of the SPL at every frequency band. Its calculation is based on the assumption that the sound signal is sinusoidally amplitude-modulated. The calculation procedure for this metric is described in Fig. 4 of the Lee *et al.* study,[13] and the same procedure was used in the present study. A more detailed description of this calculation can be found in the Appendix.

2. Stimulus generation

Eight WTN samples were extracted from four original recordings (samples 1–8). Sample pairs 1–2, 3–4, 5–6, and 7–8 were extracted from the same recording, and thus contained similar background noise. Four different amounts of unmodulated WTN were added to each sample, leading to 32 stimuli (4 AM depths $\times$ 8 samples) with reduced MD. The unmodulated noise level was varied so that $S_i(t) = S_o(t) + \nu\, N(t)$, where $S_i(t)$, $S_o(t)$, and $N(t)$ are the time-domain signals for the ith stimulus, the original sample, and the added unmodulated WTN, respectively, and ν is a weighting factor for the amount of added unmodulated WTN that was set to either 0.5, 0.7, 1, or 2. In addition to these 32 stimuli, the original samples without the unmodulated WTN were also used as stimuli, leading to 40 stimuli in total.

3. Stimulus metrics

The spectrograms of the original samples were obtained to analyze the spectral content of the samples and the approximate frequency regions in which AM occurred. All samples presented significant low-frequency content below 200 Hz, corresponding to a combination of the steady WTN and the background noise. The spectrogram of original sample 1 is given in Fig. 2 (top panel). The frequency regions

J. Acoust. Soc. Am. **139** (3), March 2016
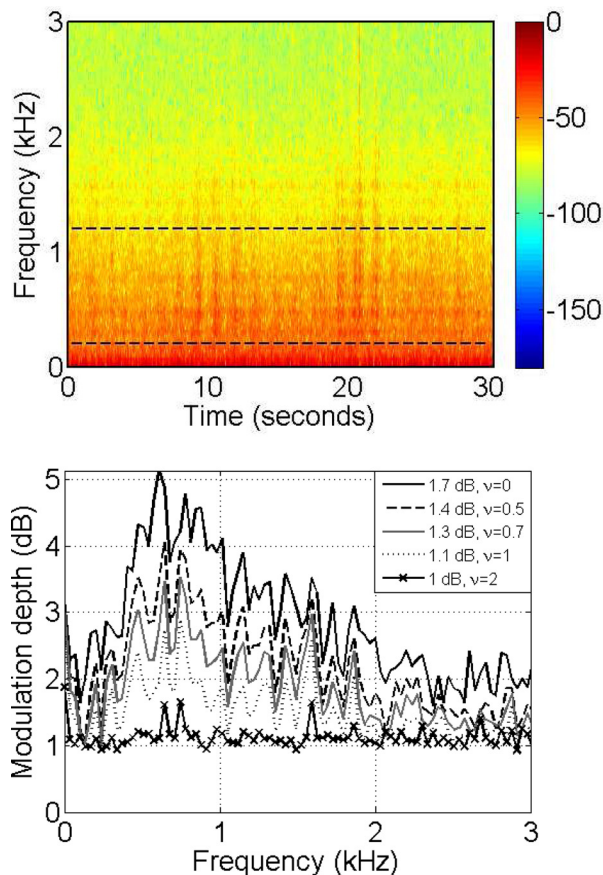
Ioannidou *et al.* 1243

FIG. 2. (Color online) (Top panel) Spectrogram of original WTN sample 1. The color scale on the right represents a dB scale with unspecified reference level. The dashed lines at 200 and 1200 Hz indicate the frequency range of the most prominent modulations. (Bottom panel) MDS for stimuli obtained from sample 1. The MDS of the original sample (black solid line) has the highest MD values of all stimuli for the whole frequency range. As unmodulated noise is added to the original sample, the MD values of the derived samples with $\nu = 0.5$ (black dashed line), $\nu = 0.7$ (grey solid line), $\nu = 1$ (black dotted line), and $\nu = 2$ (black solid line with $\times$ marker) are progressively reduced. The legend also includes the mean MD values for each stimulus.

where modulations occurred for each sample, derived from visual inspection of their corresponding spectrograms, are summarized in Table II.

The modulation depth spectra (MDS) for sample 1 with different amounts of added unmodulated WTN can be seen in Fig. 2 (bottom panel). The numbers shown in the legend are

the mean MD values across the spectrum for each stimulus. Figure 2 (bottom panel) depicts how the MD is reduced when unmodulated WTN is added to the original stimulus. The mean MD values for all samples can be found in Table III.

Although the MDS represents the difference between the minimum and maximum SPL for different octave bands, it does not reflect the absolute value of the SPL range in each band. For example, a sound may present high MD values at octave bands with relatively low SPL, which are not perceived by a listener. Therefore, the MDS curves should be considered together with the spectrogram used for their calculation.

This study suggests another way for quantifying the modulation depth, i.e., the relative modulation strength. For this calculation the mean MDS value of each stimulus, as shown in the legend of Fig. 2 (bottom) is divided by the highest mean MDS value, which corresponds to the original sample. The result is a percentage indicating the strength of modulation. The calculation is done in such a way that the original sample corresponds to 100% of modulation, while the stimulus with the lowest MDS value corresponds to 0% modulation.

4. Statistical analysis

The mean annoyance ratings for each listener and condition over all repetitions were calculated. The means and standard deviations (SDs) of annoyance ratings across listeners were then obtained. A two-way analysis of variance (ANOVA) was performed with mean MD values and sample as factors and interaction included in the model. The modulation strength of each processed sample relative to that of the corresponding fully modulated original sample was calculated and a two-way ANOVA with interaction was also performed with this relative modulation strength and sample as factors. Moreover, a one-way ANOVA was also performed to compare the relative modulation strength of the eight original samples. A significance level of 0.05 was used. The data were normally distributed, which justifies the use of means, SD, and ANOVA.

B. Results and discussion

The mean and 95% confidence intervals of the annoyance ratings over 10 subjects are presented in Fig. 3 as a

TABLE II. For each WTN sample, approximate distance of outdoor measurement locations from wind turbines, number of wind turbines, total capacity of wind turbines, and AM spectral regions derived from visual inspection of the spectrograms of samples 1 to 8. The AM frequency, MD(t) range, selected κ, λ values in pairs, and resulting AM type of samples 1, 3, 5, and 7 are also shown. For these samples, the AM frequencies are estimated according to the method presented in Ref. 16. κ and λ are parameters that were used to control the levels of the unmodulated WTN wtn(t) and the background noise, respectively.

Sample	WT Distance (m)	WT Number/Total capacity	Spectral region (Hz)	AM frequency (Hz)	MD range (dB)	κ, λ	AM type
1	≈700	8 (16 MW)	200–1200	0.69	5–12	0.3, 0.4	Periods of OAM
2	≈700	8 (16 MW)	200–800	—			
3	≈400	4 (9.2 MW)	200–600	0.76	3–8	0.2, 0.1	Periods of OAM
4	≈400	4 (9.2 MW)	250–800	—			
5	Unknown	3 (9 MW)	200–1200	0.87	<6	0.4, 0.01	NAM
6	Unknown	3 (9 MW)	250–600	—			
7	≈550	2 (2.05 MW)	500–1200	0.63	3–15	0.2, 0.01	Periods of OAM
8	≈550	2 (2.05 MW)	450–1200	—			

TABLE III. Mean MD values of samples 1 to 8 for the original sample and the four stimuli with added unmodulated WTN (ν set to 0.5, 0.7, 1, and 2) for the frequency range from 0 to 3 kHz.

Sample	Mean MD values (dB)							
	1	2	3	4	5	6	7	8
original	1.7	1.6	1.9	1.3	1.4	2	3.7	2.4
$\nu = 0.5$	1.4	1.4	1.5	1.1	1.1	1.5	2.4	1.9
$\nu = 0.7$	1.3	1.3	1.4	1.1	1.1	1.3	2	1.7
$\nu = 1$	1.1	1.2	1.2	1	1	1.2	1.6	1.5
$\nu = 2$	1	1	0.9	1	1	1	1.1	1.3

function of the mean MD across the spectrum for each of the original samples and the four corresponding stimuli with added unmodulated WTN. For each sample, the original stimulus always led to the highest annoyance. For all stimuli, annoyance decreased as the mean MD values were reduced compared to the original stimulus, in accordance with the results from earlier studies.[13]

Both mean MD value [F (1,384) = 1675.98, p < 0.001] and sample [F (7,384) = 28.54, p < 0.001] had a significant effect on annoyance, and there was a significant interaction between mean MD and sample [F(7,384) = 63.6, p < 0.001].

Figure 4 shows the annoyance response as a function of relative modulation strength. For each stimulus, the relative modulation strength is shown as a percentage of the mean MD value of the corresponding original stimulus. The effect of relative modulation strength on annoyance was significant [F(1,384) = 128.72, p < 0.001], while the main effect of sample [F(7,384) = 1.71, p = 0.106] and the interaction between relative modulation strength and sample [F(7,384) = 0.85, p = 0.545] were not significant. A one-way ANOVA showed that the differences in annoyance ratings between the eight original stimuli (relative modulation strength of 100%) were not significant [F (1,72) = 1.78, p = 0.105].

Overall, it can be concluded from experiment 1 that AM depth, investigated in terms of the mean MD across the

FIG. 4. (Color online) Mean annoyance ratings and 95% confidence intervals as a function of the relative modulation strength for the same eight samples as in Fig. 3.

spectrum, is a significant parameter in determining WTN annoyance when the overall stimulus level is fixed. This is consistent with the findings of Lee et al.,[13] the only difference between these studies being that far-field stimuli were used here, while Lee et al. used near-field stimuli for the listening experiments. Thus, while the mean MD appears to be essential for perceived annoyance, the significant effect of sample and its interaction with mean MD, as well as the lack of significant difference in annoyance between the original samples despite different mean MDs, suggest that other attributes of the samples than overall level or AM depth also affect annoyance. Such attributes could be related to the spectral content of the WTN, AM frequency, or the intermittence of the AM phenomenon. The following experiment investigated a possible role of AM frequency in WTN annoyance.

IV. EXPERIMENT 2: EFFECT OF AM FREQUENCY ON ANNOYANCE

Modulation frequency is a parameter that has been found to have a significant effect on annoyance when using stimuli with constant modulation depth over time.[2] However, in original WTN samples, the modulation depth is not constant but typically varies over time. Moreover, the MDS is an indication of the AM strength in the frequency domain but does not account for these WTN temporal variations that may be important for perceived annoyance.

Therefore, the effect of AM frequency on annoyance was investigated here using time-varying amplitude-modulated WTN samples that were synthesized by generating steady amplitude-modulated stimuli with constant amplitude as in a recent study.[14] This made it possible to generate synthesized WTN stimuli with controlled AM parameters based on the original WTN samples. The present synthesis method is more realistic as it takes into account the temporal variation of AM depth and the intermittence of the AM phenomenon (mostly occurring for OAM). In experiment 2, synthesized stimuli with typical modulation frequencies for
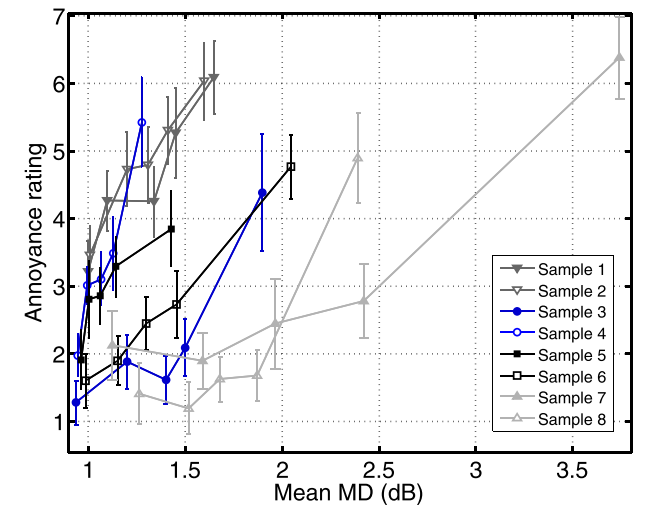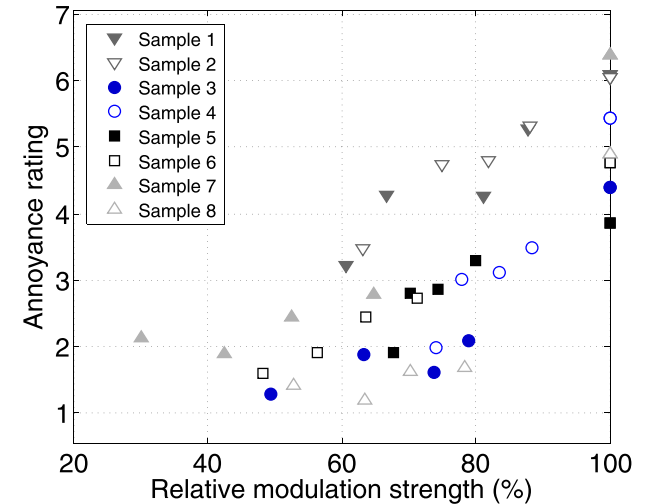

FIG. 3. (Color online) Mean annoyance ratings and 95% confidence intervals as a function of the mean MD value for samples 1–2 (grey, down triangles), 3–4 (blue, circles), 5–6 (black, squares), and 7–8 (light grey, up triangles). As samples in each pair were extracted from the same recording, the same marker shape is used per pair.

modern wind turbines were generated to investigate whether AM frequency is a statistically significant parameter in determining annoyance.

A. Method

1. Amplitude-modulated noise synthesis

A discrete time model for an amplitude-modulated WTN signal[14] can be expressed as $x(n) = w(n) [1 + \mu \cos(2\pi f_m n)]$, where $x(n)$ is the acoustic pressure of the signal at time sample n, $w(n)$ is a Gaussian white noise process with variance σ^2, μ is the modulation index, and f_m is the modulation frequency.

Based on this expression, a simplified equation for the generation of an amplitude-modulated WTN stimulus from an original WTN sample was used in the present study:

$$x(t) = A_c \text{wtn}(t)[1 + \mu \cos(2\pi f_m t)], \tag{1}$$

where $\text{wtn}(t)$ is the carrier broadband signal, i.e., low-pass-filtered Gaussian white noise (see Sec. III A), and A_c is an amplitude factor. To extend the existing synthesis method, this equation was modified as follows:

(1) Instead of a constant amplitude, A_c, and modulation index, μ, a fluctuating amplitude over time, $A_{env}(t)$, and a new AM metric, the modulation index over time $\mu(t)$, were used to introduce the modulation fluctuation.
(2) Modulated WTN normally consists of two parts: a steady noise source and a modulated noise source.[14] Here, the unmodulated $\text{wtn}(t)$ was band-pass-filtered between 200 and 1200 Hz, i.e., the region in which AM (NAM or OAM) is the strongest, leading to a band-pass unmodulated noise, $\text{wtn}_{bp}(t)$. The modulation was then applied only to $\text{wtn}_{bp}(t)$ and the steady noise was added to the amplitude-modulated noise.
(3) The background noise of the original WTN sample, derived from its band-pass-filtering between 2 and 10 kHz, was added to the amplitude-modulated synthesized noise. This resulted in stimuli that had the same background noise as the original samples and thus sounded more realistic.

The final expression for the amplitude-modulated WTN was

$$x(t) = \{A_{env}(t)\text{wtn}_{bp}(t) [1 + \mu(t) \cdot \cos(2\pi f_m t)]\} + \kappa \cdot \text{wtn}(t) + \lambda \cdot \text{bkn}(t), \tag{2}$$

where $A_{env}(t)$ is the envelope of the modulated signal, $\text{wtn}_{bp}(t)$ the unmodulated band-pass WTN, $\mu(t)$ the modulation index over time, $\text{wtn}(t)$ the unmodulated WTN, $\text{bkn}(t)$ the background noise from the sample, and $\kappa, \lambda \in R > 0$ are parameters that control the levels of the unmodulated WTN and the background noise of the original sample, respectively. The first term of the equation denotes the amplitude-modulated sound.

The assumptions that were made during the synthesis procedure led to certain limitations. First of all, a sinusoidally amplitude-modulated WTN was assumed and the

values of f_m, $\mu(t)$, and modulation depth (MD) were estimated based on this assumption. This led to approximations of the original WTN recordings, as WTN is not truly sinusoidally amplitude-modulated.[14] Moreover, the same filter was applied to all original WTN samples to determine their background noise, while these samples were recorded under different conditions and may thus have had different background-noise characteristics. Finally, the parameters κ and λ were changed manually in order to generate stimuli with similar spectral shape to the original WTN samples. Due to these limitations, it was not possible to generate stimuli with the exact same $\mu(t)$ and MD as the original WTN samples. However, this synthesis procedure was satisfactory for the present purpose to generate stimuli that sounded realistic while containing intermittent AM with MD that varied over time and AM parameters that could be systematically varied in a controlled way.

2. Estimation of AM parameters

Original WTN samples 1, 3, 5, and 7 from experiment 1 were processed to estimate the AM parameters used for WTN synthesis. As the method used for this estimation is more efficient for clearly discernible modulations, these samples were band-pass filtered between 200 and 1200 Hz prior to their analysis to focus on the octave band where the AM was the strongest and to avoid out-of-band noise.

Initially, the fluctuating amplitude over time $A_{env}(t)$ was introduced using the envelope of each original WTN sample. Then, to obtain $\mu(t)$ and MD estimates, it was crucial to first estimate the modulation frequency f_m. Assuming that WTN is sinusoidally amplitude-modulated, a maximum likelihood estimator of its f_m can be derived by locating the maximum value of the Fourier transform of the squared signal, an approach applied in the field of WTN,[14] known as short-time energy analysis. The Fourier transform of the squared signal yields the magnitude of the short-time energy over modulation frequency. This procedure is explained in Fig. 7 of Ref. 14 and was used in this study. When the envelope variations of a sound are quantified, the MD can be calculated for each time slot corresponding to one period $1/f_m$ of the WTN AM. As WTN is a quasi-periodic noise, these time slots are approximately equal to each other, as the period varies only slightly over time. Taking the example of the signal envelope of original sample 1 (Fig. 5), the consecutive peak-trough pairs corresponding to a given time slot were first identified (grey triangles). The modulation index μ for each time slot was calculated as $\mu = (A_{peak} - A_{trough})/(A_{peak} + A_{trough})$, where A_{peak} and A_{trough} are the peak and trough amplitudes for this particular time slot, respectively. The μ values for all consecutive time slots resulted in a time-dependent $\mu(t)$ for original sample 1, which can be seen in Fig. 5 (black line). The modulation depth MD for each time slot was defined as $MD = 20\log_{10} (1+\mu)/(1-\mu) = 20 \log_{10}(A_{peak}/A_{trough})$ (dB). The MD values for all consecutive time slots resulted in the modulation depth over time $MD(t)$ of original sample 1, illustrated in Fig. 6(a) (grey line).
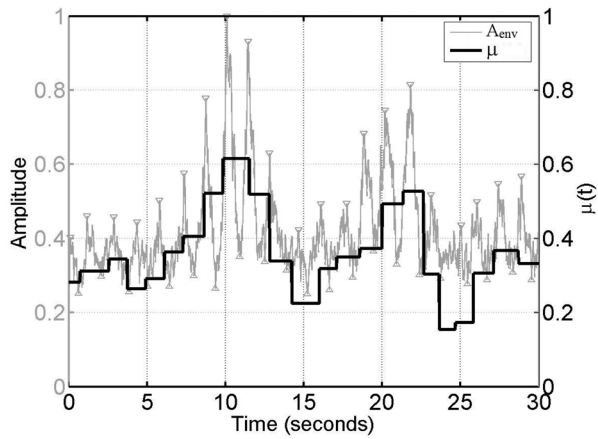
FIG. 5. Amplitude of signal envelope of original WTN sample 1 (left ordinate axis, grey line). The peaks and troughs are marked with down and up triangles, respectively. The corresponding modulation index μ for the selected time slot is given on the right ordinate axis (black line).

3. Stimulus generation

For each of the four original WTN samples, the estimated $A_{env}(t)$, $\mu(t)$, and f_m were inserted into Eq. (2) and the unmodulated WTN, wtn(t), was obtained as described in Sec. III A. Values for κ and λ were selected manually such that the generated stimuli sounded realistic and had similar spectral shapes to those of the corresponding original WTN samples. The four generated broadband WTN stimuli were finally low-pass filtered with a cutoff frequency of 6.5 kHz.

Figures 6(a)–6(d) show the paired comparisons between the MD(t) of original WTN samples 1, 3, 5, and 7 (grey lines) and the corresponding synthesized stimuli (black lines). Due to the limitations of the synthesis method (see Sec. IV A 1), the synthesized stimuli presented deviations from the MD(t) values of the original WTN samples. The synthesis was most accurate for sample 1, as the generated stimulus succeeded in reproducing the main $\mu(t)$ peaks of the original WTN sample. For sample 3, the synthesis resulted in overestimated $\mu(t)$ values. For sample 5, it identified only few peaks. For sample 7, it failed to represent fluctuations in the first 5 s of the sample.

Despite the above discrepancies, the aim was for the synthesized stimuli to present intermittent AM with realistic characteristics and to sound natural such that they were not crucial for the present purpose. In order to verify that the synthesized stimuli were perceptually similar to original WTN, an additional listening test was carried out, in which the participants rated the annoyance caused by the original samples and the corresponding synthesized stimuli. The paired comparisons between the annoyance ratings for original samples and their corresponding synthesized stimuli were always statistically insignificant. Moreover, when asked at the end of the experiment, the participants reported that all samples sounded realistic to them and that they were not able to tell which samples were synthesized and which were original. Therefore, the synthesis accuracy was deemed sufficient to provide realistic-sounding WTN.

Table II shows the obtained modulation frequency estimates f_m and the MD(t) range of the original WTN samples. This table also shows the selected κ, λ values for each one of the generated stimuli. The MD(t) range of the original WTN
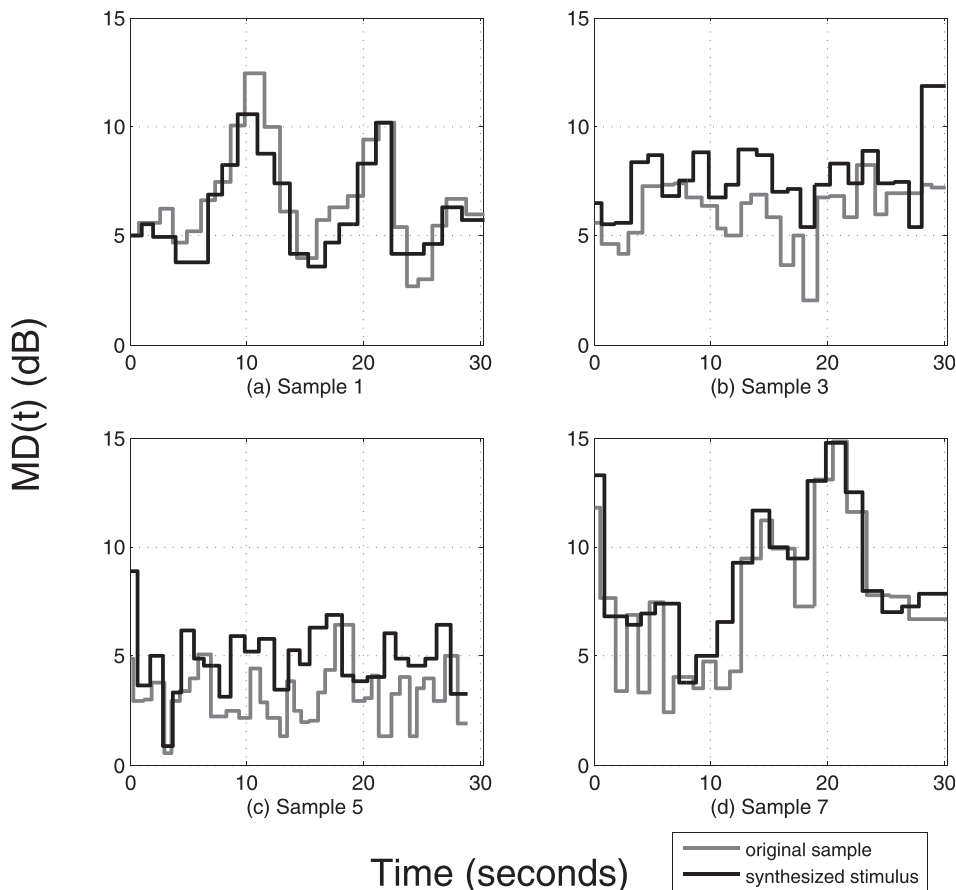

FIG. 6. (a)–(d) AM depth over time for original WTN samples 1, 3, 5, and 7 (grey line) and their corresponding synthesized stimuli (black line).

samples combined with the frequency range within which modulations occur were used to evaluate whether a signal was dominated by NAM or included periods of OAM. Stimuli that presented MD(t) values greater than 6 dB predominantly at frequencies between 300 and 400 Hz, were considered to present OAM periods. Stimuli that presented MD(t) values below 6 dB mainly at frequencies between 400 and 1000 Hz were considered to be dominated by NAM periods. Therefore, original stimuli 1, 3, and 7, with modulation depths above 6 dB, were considered to be presenting OAM periods. For original stimulus 5, modulation depths smaller than 6 dB implied a dominance of NAM.

Finally, as AM frequency was the parameter of interest in experiment 2, the f_m of each synthesized stimulus, originally set to that of its corresponding WTN sample, was artificially fixed to either 0.5, 1, 1.5, or 2 Hz, leading to four synthesized stimuli for each original WTN sample, i.e., 16 stimuli in total. Values between 0.5 and 2 Hz are considered to be typical f_m values for most modern large-scale wind turbines.[2]

4. Statistical analysis

The mean annoyance ratings for each listener and condition over all repetitions were calculated. The means and standard deviations of annoyance ratings across listeners were then obtained. A two-way ANOVA was performed with f_m and sample as factors and interaction included in the model. A significance level of 0.05 was used.

B. Results and discussion

The mean and 95% confidence intervals of the annoyance ratings over 14 subjects for each of the 16 stimuli are given in Fig. 7 as a function of f_m. Overall, an increase in modulation frequency resulted in a rise in annoyance. Despite this clear trend, the main effect of f_m was insignificant [F (3,208) = 2.56, p = 0.056], while the effect of sample was significant [F (3,208) = 11.89, p < 0.001]. There was no
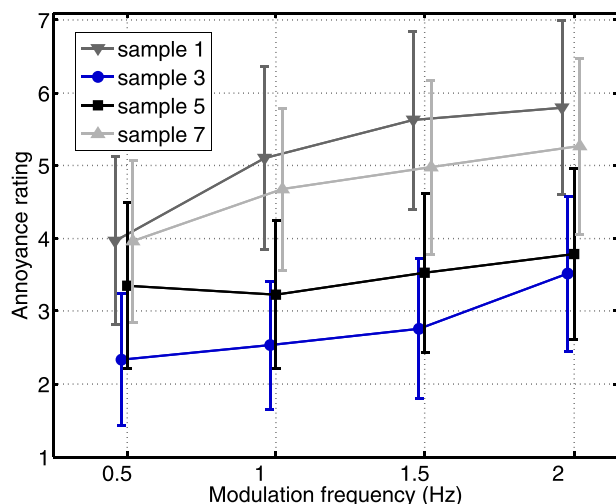

FIG. 7. (Color online) Mean annoyance ratings and 95% confidence intervals as a function of modulation frequency f_m for synthesized stimuli based on samples 1 (grey, down triangles), 3 (blue, circles), 5 (black, squares), and 7 (light grey, up triangles).

significant interaction between f_m and sample [F (9,208) = 0.26, p = 0.985].

Overall, AM frequency was found to have only a limited effect on WTN annoyance when the temporal variations in AM strength are taken into account. This contrasts with earlier findings obtained with static AM strength stimuli.[2] The fact that the effect was only borderline significant may be partly because the present experiment took temporal AM variations into account.

It is also worth noting that the effect of AM frequency was slightly more pronounced for the two samples with overall higher annoyance ratings (Fig. 7, grey triangles) than for the other two samples (blue circles and black squares). Therefore, the importance of AM frequency for annoyance may interact with other attributes that affect annoyance, such as overall AM strength or AM intermittence. The following experiment focused on the effect of the presence of intermittent periods of OAM on WTN annoyance.

V. EXPERIMENT 3: EFFECT OF AM INTERMITTENCE ON ANNOYANCE

The annoyance of steady WTN AM sounds was recently found not to significantly depend on the AM type (NAM or OAM), once the A-weighted sound pressure level was taken into account.[2] However, previous studies did not account for temporal variations of AM. Here, synthesized stimuli containing NAM or OAM and accounting for these temporal variations were developed by modifying the synthesis method used in experiment 2, to generate stimuli that were either dominated by NAM or included intermittent OAM periods. The AM metrics derived using the methodology of experiment 3 were more precise than the MDS metric of experiment 1 in the sense that they yielded temporal information and allowed a more accurate prediction of whether a signal included only NAM or also intermittent periods of OAM. The aim of experiment 3 was to investigate the effect of AM type on annoyance and to determine whether AM depth affects annoyance differently depending on AM type, using synthesized NAM and OAM stimuli with various MD(t) values.

A. Method

1. Stimulus generation

The typical spectral regions, in which the most prominent NAM and OAM modulations occur according to literature (see Table I), differ from the AM spectral regions of the original WTN samples of the present study (see Table II). This can be explained by the fact that while a sample may present its most prominent NAM or OAM within a limited frequency range, the actual modulations may cover a wider frequency range. Therefore, the band-pass filtering of wtn$_{bp}$(t) for NAM/OAM synthesis focused on two frequency regions, one between 600 and 1200 Hz for NAM and the other between 200 and 400 Hz for OAM. These filters were also selected to correspond to the upper and lower filtering limits used in experiment 2 (200–1200 Hz).

The NAM modulation depth typically varies around a constant value below 6 dB, such that a constant number μ_{NAM} was selected to the modulation index in the NAM model. NAM is a continuous characteristic of amplitude-modulated WTN. Thus, the amplitude of the envelope, $A_{env}(t)$, was set to 1 so as not to introduce level fluctuations. According to Eq. (2), the expression for the NAM stimuli was $x_{NAM}(t) = \text{wtn}_{bp-mid}(t) \cdot [1 + \mu_{NAM} \cos(2\pi f_m t)] + \kappa \, \text{wtn}(t) + \lambda \, \text{bkn}(t)$, where $\text{wtn}_{bp-mid}(t)$ is the unmodulated band-pass WTN filtered for NAM.

The OAM modulation depth typically fluctuates between 6 and 12 dB, and therefore the $\mu(t)$ value from the original WTN samples was chosen as the modulation index in the OAM model. OAM is an intermittent phenomenon, and the signal envelope, $A_{env}(t)$, of the original WTN samples was used to introduce level fluctuations. From Eq. (1), the expression for the OAM stimuli was $x_{OAM}(t) = \{A_{env}(t) \cdot \text{wtn}_{bp-low}(t)[1 + \mu(t) \cdot \cos(2\pi f_m t)]\} + \kappa \, \text{wtn}(t) + \lambda \, \text{bkn}(t)$, where $\text{wtn}_{bp-low}(t)$ is the unmodulated band-pass WTN filtered for OAM.

The NAM/OAM synthesis was based on the parameter estimates from original WTN sample 1 to generate stimuli with the same unmodulated WTN and background noise. To create stimuli with various $MD(t)$ values corresponding to typically observed modulation depths for NAM/OAM, different $\mu(t)$ values were inserted into the NAM/OAM equations. The values for κ and λ were kept constant for all generated stimuli, with values of 2.5 and 0.2, respectively.

Three NAM stimuli were generated for μ_{NAM} set to 0.2, 0.5, and 1, resulting in mean MD values of 3.5, 4.5, and 5.2 dB, respectively (Fig. 8, top panel).

The OAM stimulus generation was based on the intermittence pattern of WTN sample 1, which $\mu(t)$ (Fig. 5, black line) and thus $MD(t)$ [Fig. 6(a), grey line] presented maxima around 10 and 20 s. In order to obtain OAM stimuli with various $MD(t)$ values, these peak values were modified, resulting in the stimuli with intermittent AM and MD values within the OAM modulation depth range (6–12 dB).

From the expression for modulation depth, $MD = 20 \log_{10}(1+\mu)/(1-\mu)$, a modulation depth greater than 6 dB corresponds to a modulation index $\mu > 0.33$. Here, the $\mu(t)$ peak values larger than 0.33 in Fig. 5 (black line) were modified, resulting in three $\mu(t)$ values which were inserted into the model, leading to three OAM stimuli with mean MD values of 5.6, 6.5, and 7 dB (Fig. 8, bottom panel).

2. Statistical analysis

The mean annoyance ratings for each listener and condition over all repetitions were calculated. The means and standard deviations of annoyance ratings across listeners were then obtained. A two-way ANOVA was performed with mean MD values and AM type as factors and interaction included in the model. A significance level of 0.05 was used.

B. Results and discussion

The mean annoyance ratings over 13 subjects for each of the 6 stimuli are given as a function of their mean MD
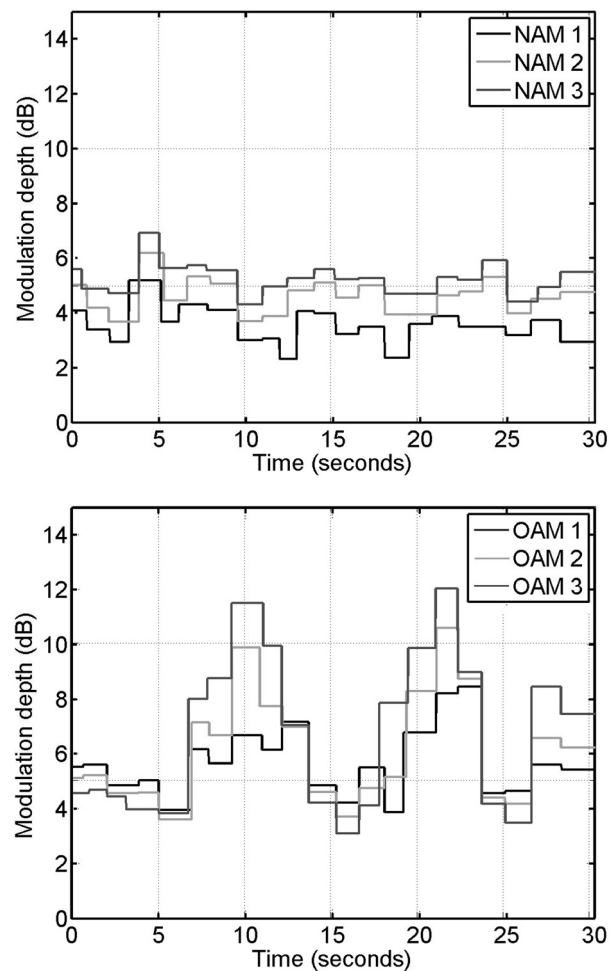

FIG. 8. (Top panel) MD(t) of the generated stimuli NAM 1 (black line), 2 (light grey line), and 3 (grey line) with mean MD values of 3.5, 4.5, and 5.2 dB, respectively. (Bottom panel) MD(t) of the generated stimuli OAM 1 (black line), 2 (light grey line), and 3 (grey line) with mean MD values of 5.6, 6.5, and 7 dB, respectively.

value in Fig. 9. Error bars indicate 95% confidence intervals. While annoyance ratings increased with mean MD for both NAM and OAM stimuli independently, confirming the effect of AM depth on annoyance found in experiment 1, the overall annoyance response was not higher for the OAM than for the NAM stimuli. While mean MD had a significant effect on annoyance [$F_{(1,74)} = 4.72$, $p = 0.033$], AM type did not [$F_{(1,74)} = 0.01$, $p = 0.919$]. There was no interaction between the two factors [$F_{(1,74)} = 0.21$, $p = 0.652$].

This suggests that the "baseline" MD outside OAM periods, which was similar between the NAM and OAM stimuli used here, is what mainly determines annoyance, and that deviations from this baseline in the form of intermittent OAM periods do not increase annoyance further. Therefore, for a fixed presentation level, the annoyance caused by WTN might not be related to the frequency region in which the modulation is most prominent, as this region typically differs for NAM and OAM. This conclusion is in line with the outcome of a recent study, which stated that "the response to the noise was not significantly affected by the frequency content of the modulated noise (NAM or OAM), once the A-weighted sound pressure level was taken into account."[2]

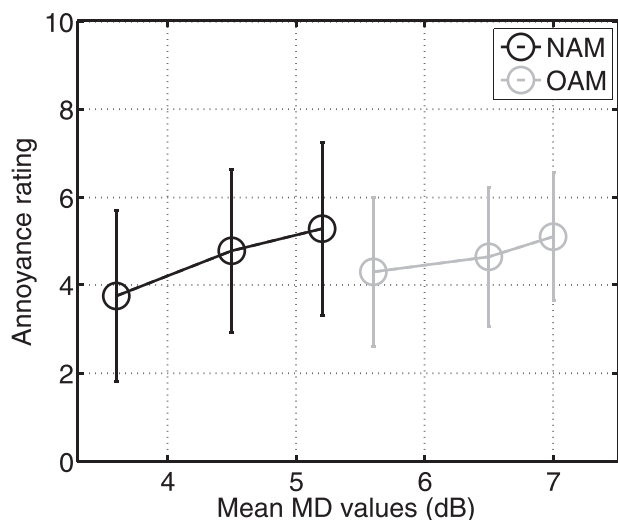J. Acoust. Soc. Am. **139** (3), March 2016

Ioannidou *et al.* 1249

FIG. 9. Mean annoyance ratings and 95% confidence intervals as a function of the mean MD values for NAM (black line, circles) and OAM (grey line, triangles) stimuli.

Here, it was found that, when the temporal variations in MD are taken into account in the synthesized WTN stimuli, intermittence periods of OAM do not lead to increased annoyance.

VI. CONCLUSION

Listening experiments were conducted in a controlled laboratory setting to investigate the effect of AM depth, frequency, and type (NAM only vs intermittent OAM periods) on WTN annoyance. Far-field original WTN samples were used to synthesize realistic-sounding stimuli in which these parameters could be varied systematically, taking into account the intermittence and temporal variations of the AM phenomenon as they occur in the field. Experiment 1 showed that, for a fixed presentation level, the mean AM depth was a significant parameter in determining annoyance, thus extending previous findings obtained with near-field stimuli to the far field. In experiments 2 and 3, time-varying AM depth was introduced in the stimuli to mimic the real temporal AM fluctuations of WTN. The results of experiment 2 showed that, for a fixed presentation level, AM frequency had only a limited effect on perceived annoyance, which was negligible for samples with low overall annoyance based on other factors. In experiment 3, it was found that the annoyance caused by stimuli that were dominated by NAM or included intermittent OAM periods did not differ when their mean "baseline" AM depth outside OAM periods was similar. Therefore, the frequency regions in which the modulations are most prominent and the presence of intermittent OAM periods might not be as crucial for perceived annoyance as the baseline AM depth of the noise in NAM-dominated periods. This suggests that this latter parameter is the one that ought to be reduced to have the most beneficial impact on WTN annoyance.

Finally, it should be noted that the present findings are only valid for the present controlled listening conditions, i.e., a short listening time in an artificial laboratory setting. The limited number of listeners limited the statistical power

necessary for strong conclusions on the effect of AM modulation, frequency, and type on annoyance. It would thus be beneficial to investigate whether their validity can be extended to considerably longer exposure times in a field setting, where the intermittent nature of OAM may have a different impact.

APPENDIX

A more detailed explanation of how MDS is calculated is presented below. The calculation is done according to the method used in the Lee et al. study.[14]

The spectrogram of the original sample was calculated by applying the Fourier transform to each time step of it. For this calculation, it was important to define the frequency and time resolution of the spectrogram so as to resolve the amplitude modulation (for all spectrograms: $t_{res} = 1\,ms$ and $f_{res} = 65\,Hz$). The time domain of original sample 1 can be seen in Fig. 10.

The spectrogram of original sample 1, shown in Fig. 2 (top panel), illustrates the magnitude of the signal for each frequency band and time slots. Two regions with discernible modulation peaks can be seen around 10 and 20 s, occurring for frequencies mainly below 2 kHz.

Applying a second Fourier transform to each frequency band of the spectrogram results in the Fourier transform of the magnitude over time. For the frequency band around 1 kHz, this resulting signal can be seen in Fig. 11.

Figure 11 presents two dominant peaks (red lines): one at 0 Hz, which corresponds to the steady root mean square value of the signal, p_1, and one at a frequency which is estimated to be the modulation frequency $f_m = 0.69\,Hz$ and corresponds to the sinusoidal amplitude modulation of the
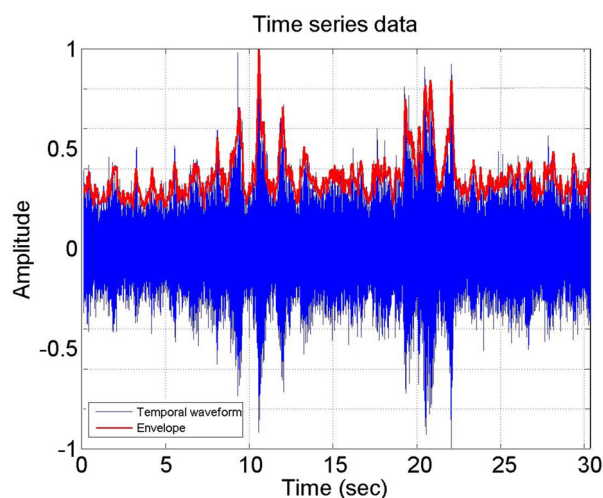

FIG. 10. (Color online) Time domain of original sample 1 (blue line). The envelope of the signal is highlighted in red.
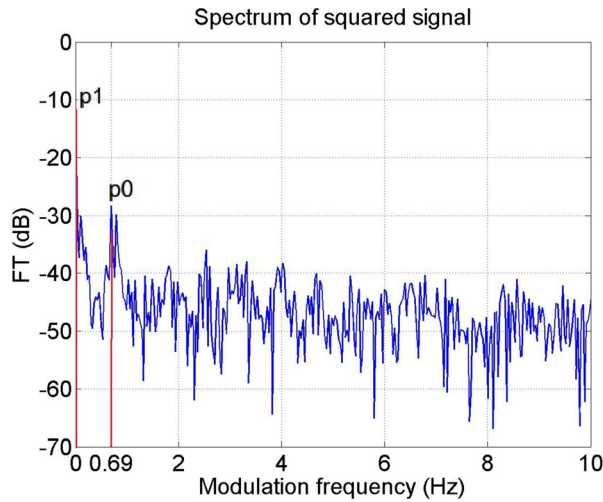
FIG. 11. (Color online) The Fourier transform of the magnitude of original sample 1 (blue line). The peak at 0 Hz (red line) represents the steady root mean square value of the signal, p_1. The peak at 0.69 Hz represents the sinusoidal amplitude modulation of the signal, p_0. The frequency $f_m = 0.69$ Hz is estimated to be the modulation frequency.

signal, p_0. The selection of 1 ms as the time resolution of the spectrogram leads to the clear identification of the peak which corresponds to the modulation frequency.

Assuming that WTN is sinusoidally amplitude modulated, all other values except these two peaks can be neglected. Then, applying the inverse Fourier transform to this simplified two-peak signal yields a sinusoidal signal of p_{rms} sound pressure over time. This signal is the equivalent of the term $[1 + \mu \cdot \cos(2\pi f_m n)]$ in the equation of the discrete time model for amplitude-modulated WTN [Eq. (1)].[15] The MD for the frequency band around 1 kHz is defined as the difference between the maximum and minimum values of the sound pressure of this signal: $MD = 20 \log[(p_1 + p_0)/(p_1 - p_0)] = 20 \log(p_{max}/p_{min})$ (dB). Applying this equation to all specified frequency octaves leads to the calculation of the MDS, which is given in Fig. 2 (bottom, black line) for original sample 1.

[1]R. Meir, L. Legerton, M. B. Anderson, B. Berry, A. Bullmore, M. Hayes, M. Jiggins, E. Leeming, P. Musgrove, D. J. Spode, H. A. Thomas, E. Tomalin, M. Trinick, and J. Warren, "The assessment and rating of noise from wind farms. ETSU-R-97," Technical report (Department of Trade and Industry, UK, 1996).

[2]S. von Huenerbein, B. Piper, M. Cand, and A. King, "Development of an AM Dose-Response Relationship," Work Package B2, University of Salford, Acoustics Research Centre (2012).

[3]S. Oerlemans, "An explanation for enhanced amplitude modulation of wind turbine noise," Technical report No NLR-CR-2011-071 (National Aerospace Laboratory, N. L. R., The Netherlands, 2011).

[4]K. Kim, S. Lee, S. Lee, and H. Kim, "Perception of amplitude-modulated noise and from wind turbines," in *17th International Congress on Sound & Vibration*, Cairo, Egypt (2010).

[5]G. P. van den Berg, "Effect of the wind profile at night on wind turbine sound," J. Sound Vib. **277**, 955–970 (2004).

[6]A. J. Bullmore, J. F. Lowson, J. H. Bass, and P. Dunbabin, "Wind turbine measurements for noise source identification," Technical report, ETSU W/13/00391/00/REP (Department of Trade and Industry, UK, 1999).

[7]G. P. van den Berg, "Do wind turbines produce significant low frequency sound levels?," in *11th International Meeting on Low Frequency Noise and Vibration*, Maastricht, The Netherlands (2004).

[8]M. G. Smith, "Wind turbine amplitude modulation: Research to improve understanding as to its cause and effect," Consultancy report, Ref: 8630-R01, University of Southampton (2012).

[9]E. Zwicker, "A proposal for defining and calculating the unbiased annoyance," Z. Laermbekaempfung (Mag. Noise Control) **38**, 91–97 (1991).

[10]C. Kantarelis and J. G. Walker, "The identification and subjective effect of amplitude modulation in diesel engine exhaust noise," J. Sound Vib. **120**, 297–302 (1988).

[11]F. van der Ploeg, E. Buikema, J. Vos, and M. M. J. Houben, "Annoyance caused by low frequency sounds—spectral and temporal effects," in *Internoise 2010*, Lisbon, Portugal (2010).

[12]S. Lee, H. Kim, S. Lee, and K. Kim, "An estimation method of the amplitude modulation in wind turbine noise for community response assessment," in *Third International Meeting on Wind Turbine Noise*, Aalborg, Denmark (2009).

[13]S. Lee, K. Kim, W. Choi, and S. Lee, "Annoyance caused by amplitude modulation of wind turbine noise," Noise Control Eng. J. **59**, 38–46 (2011).

[14]P. White, "The measurement and definition of amplitude modulations," Technical report (University of Southampton, Institute of Sound and Vibration, UK, 2012).

[15]T. Poulsen, "Influence of session length on judged annoyance," J. Sound Vib. **145**(2), 217–224 (1991).

[16]H. Møller, "Annoyance of audible infrasound," J. Low Freq. Noise Vib. **6**(1) 1–17 (1987), available at: http://vbn.aau.dk/ws/files/227874746/1987_M_ller_LF_journal.pdf.