



**COVER PAGE**

***Document downloaded by @DAEL***

***Sat May 9 14:36:39 2026***

***For personal use***

When automatic English translation is provided, only the original document is authentic.

The EAA cannot be held responsible of any translation error

Bibliographical reference

*Annoyance and Partial Masking of Wind Turbine Noise from Ambient Sources*, Anders Johansson, Karl Bolin and Jesper Alvarsson, *Acta Acustica* **vol. 105** (Number 6), 2019, pp. 1035-1041

DOI

<https://doi.org/10.3813/AAA.919382>

# Annoyance and Partial Masking of Wind Turbine Noise from Ambient Sources

Anders Johansson<sup>1)</sup>, Karl Bolin<sup>1)</sup>, Jesper Alvarsson<sup>1,2)</sup>

<sup>1)</sup> The Marcus Wallenberg Laboratory, Department of Aeronautical and Vehicle Engineering, Kungliga Tekniska Högskolan, Teknikringen 8, 100 44 Stockholm, Sweden. aebjo@kth.se

<sup>2)</sup> Department of Social Sciences, Södertörn University, Stockholm, Sweden

*Colour Figures: Figures in colour are given in the online version*

## Summary

This paper investigates noise annoyance from wind turbines of different sizes and in different acoustic surroundings. A listening test was conducted where wind turbine noises were rated alone and together with background sounds from a deciduous forest, a busy city and road traffic. A magnitude production procedure was implemented which showed high correlation between repeated measurements and the results were analysed using A-weighted sound levels, signal-to-noise ratios and time varying loudness and partial loudness. Ratings for wind turbine sound heard alone showed no coherent statistically significant differences between wind turbine types, neither for A-weighted sound levels nor loudness. The masking test indicate that road traffic noise is a superior masker compared to forest sound. However, these effects were only statistically significant at low sound levels, below the range 35–45 dB(A), where noise guidelines for wind turbine noise usually are stipulated.

PACS no. 43.50.-x, 43.66.-x

## 1. Introduction

Considering global warming and the Paris-agreement a continuous shift towards renewable energy is currently occurring in Europe and other parts of the world. One power source that most probably will play a significant, if not dominant, role in this transition is wind energy. In this context noise from wind turbines (WTs) is undoubtedly something that many persons will be exposed to. Epidemiological studies [21, 20, 12] have shown that sound levels around 35–40 dB(A) results in around 10% highly annoyed which is significantly lower sound levels than for transportation noise [13, 14]. The sensitivity towards wind turbine noise is also reflected in the recent WHO-environmental noise guidelines of 2018 [23].

Sound characteristics, especially a distinctive amplitude modulation, might increase the probability of detection and/or disturbance compared to steadier WT sounds and have earlier been investigated by several authors [22, 11, 24] and reviewed by Hafke-Dys *et al.* [10]. However, the perceived annoyance is not only determined by the sound itself but also by the surrounding ambient sound. It is thus likewise important to consider masking and partial masking when evaluating the environmental impact of wind turbine noise, which for instance is taken into account in the British and French wind turbine noise guidelines (ETSU-R-97 1996) [7] and (no.92-1444) [19]. Earlier investigations of masking effects of wind turbines from natural

background sounds [4, 3] have shown that both indirect and direct scaling methods yield levels of total masking around signal-to-noise (S/N) ratios of  $-10$  dB(A). In these studies, the effect of partial masking mainly occurred at negative S/N-ratios. Similar results were also observed in an epidemiological study [20], which showed that masking from road traffic noise can mitigate wind turbine noise when road sound levels are exceeding the turbine sound level by around 20 dB(A)  $L_{den}$ .

How future wind turbines will be designed and located is not yet decided, current trends favour large turbines situated in rural settings. However, there is also interest in developing smaller turbines placed in urban areas as recognized by the EU/FP7 project SWIP “New innovative solutions, components and tools for the integration of wind energy in urban and peri-urban areas”. In this perspective the question of how different wind turbine sizes are perceived in different ambient surroundings becomes an interesting question but has until now received little attention. The possibilities to explore future scenarios where wind turbine noise may be present in other surroundings than currently is naturally difficult with epidemiological surveys. However, listening tests could be used to examine future scenarios that could give an indication if it is warranted to establish different guidelines for wind turbine noises in different environments. The main objective of this paper is thus to increase our understanding of under which conditions noise annoyance from wind turbines occur and to compare if different turbine size influence the annoyance.

A-weighted sound levels have traditionally been used for noise guidelines. However, indications that loudness

Received 15 September 2019,  
accepted 5 December 2019.

might be a superior predictor of road traffic noise annoyance [18] and in the event of masked sounds it is considered probable that partial loudness levels would have the same, if not a larger, advantage. Therefore the second objective of this paper is to compare the results of the listening tests to A-weighted sound levels and to loudness and partial loudness models for time varying sounds [9] in order to evaluate if more advanced loudness models should be advocated or not.

## 2. Method

### 2.1. Equipment

#### 2.1.1. Recording equipment

The wind turbine recordings were conducted using a Bruel and Kjaer (BK) Head and Torso Simulator 4128D (HATS), with one BK 4190 microphone in each ear. The microphone signals were connected to a BK type 2690 Nexus Conditioner that was set at 1 V/Pa, high pass filter at 20 Hz and low pass filter at 22.4 kHz. The amplified signal was transmitted to a Sound Devices 788T recorder. Calibration of the microphones was done with a Larson Davis model CAL200, 1 kHz, 94 dB tone calibrator. While recording, the HATS microphones were positioned 1.5 meters above the ground. The recordings were sampled at 44.1 kHz with a bit depth of 24 bits.

#### 2.1.2. Listening test equipment

The listening test was conducted in a quiet office space and the test sounds were reproduced using Bose QC15 noise cancelling headphones. To control for a quiet listening environment, the background noise of the office was measured with a Bruel and Kjaer type 2250 sound level meter. During the measurement the microphone was placed at the same position as the ears of the listener and the integration time of the measurement was set to 2 hours (approximate duration of the test). Results showed  $L_{A,eq,2h}$  to be below 30 dB(A) in the frequency range 12.5 Hz to 20 kHz. To compensate for the non-flat frequency response of the headphones a mini-DSP board was used to equalise the sounds. This setup was considered to provide a sufficiently quiet listening environment as the headphones attenuate external noise with 20 dB above 200 Hz and has previously been used in a field study by one of the co-authors [2]. Also, magnitude production facilitates for the listeners to delay a judgement if external transient sounds is perceived but no participant reported the existence of external interference when asked after the test. Moreover, the overall high intra individual correlations coefficients, and low sound levels reported in section 3 indicate that external disturbances probably did not cause any major detriment of the judgements. The headphones and DSP-board were connected to the laptop via an Asus XONAR U7 sound-card for which the transfer amplitude at 1 kHz were independently measured and normalized for both right and left channels. Sound level adjustments were done via a Griffin Technology PowerMate rotatable knob. The test software

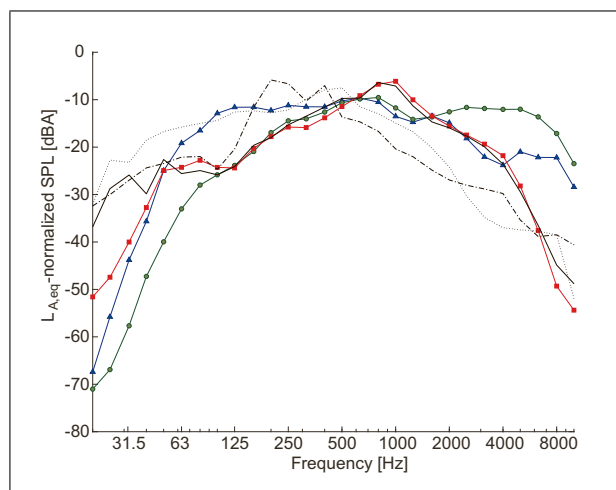


Figure 1. (Colour online) Showing the  $L_{A,eq}$ -normalized third octave spectrum for all experimental sounds. The black curves show the WT sounds, solid for small-WT, dashed-dotted for medium WT and dotted for large WT. The green (o) curve shows the woodland background, the blue ( $\Delta$ ) curve shows the city background and the red ( $\square$ ) curve represent traffic background sound.

was written in MATLAB R2014b using the GUI-interface. All post processing of measurement-data was also performed in MATLAB.

### 2.2. Sounds

#### 2.2.1. Wind turbine sounds

To represent the variation of sounds from different sized WTs, three WTs with hub-height 6, 21 and 95 meters were selected and their sounds recorded binaurally. The recordings are referred to as the small-, medium- and large-WT and third octave spectra are shown in Figure 1. Increasing low frequency content with larger turbines are observed and is in correspondence to previous literature [16]. Amplitude modulations were audible in every recording with blade passing frequencies of 3, 1 and 0.6 Hz for respective recording. It should be noted that the small-WT is of vertical axis turbine type while the two larger are of ordinary horizontal axis type. In addition, pink noise was included as a reference sound.

#### 2.2.2. Background sounds

Three background sounds were selected to represent possible WT development sites. These are recordings of (i) rustling leaves in a deciduous forest, (ii) the slow-paced sounds of pedestrians and vehicles in a city-street, (iii) a recording of traffic on a motorway. The recorded sounds are henceforth referred to as the forest-, city- and road-background noises and the third octave band spectra are shown in Figure 1.

To determine suitable levels for playback of the background sounds ten listeners, who did not participate in the main listening test, adjusted the sound level of the sounds to levels they perceived as natural. The inter-individual value of this test was near 60 dB(A), which were adopted

as the reference level for all background noises to be used in the experiment.

### 2.3. Listening test

#### 2.3.1. Test subjects

In total 34 persons, 15 women and 19 men, participated in the listening experiment. All were students or personnel at KTH at an age between 19 and 36 years and with reported normal hearing. As a gesture of gratitude for participating in the experiment the participants received cinema tickets. This group is not considered representative for people annoyed by WT noise living in rural environments which has earlier been shown by one of the co-authors to have an effect on the perceived pink-noise equivalent [2] but as the present study concerns comparisons between different sounds and environments the authors believe this misrepresentation is not diminishing the validity of the test too much.

#### 2.3.2. Scaling method

In this experiment the Borg CR100-rating scale was used together with a magnitude production procedure. The Borg CR100-scale is a level anchored ratio scale where numbers are connected to verbal anchors which describes proportions of maximal or absolute perceived sensations [5]. In this experiment the annoyance magnitudes “just noticeable”, “very weak”, “weak”, “moderate”, “strong” and “very strong” were used, which corresponds to the numbers 2–75 on a 0–100 scale, where 0 is nothing and 100 is maximum. The verbal anchors facilitate common criteria for the listeners and also permit determination of absolute levels of intensity. The Borg CR100-scale is common in exercise tests and has also been used for annoyance ratings of aircraft noise [6].

#### 2.3.3. Test procedure

During the experiment the listener was seated in front of a laptop and wore headphones for playback of the sounds. On the laptop the Borg CR100-scale was shown with an arrow pointing towards the intended magnitude. The instructions for the listener was to adjust the sound level of the stimuli until it matched the psychological magnitude of the designated level. Each listener performed four repetitions for each of the sounds and for each magnitude. The test was divided into two sessions. In the first part the listeners produced magnitudes of annoyance for the WT-sounds heard alone and in the second session background noise was included and the listeners produced magnitudes of annoyance for the part of the stimuli they perceived as WT-sound. Prior to the experiment the listeners read through test instructions including a description of the Borg CR100-scale and examples on how to interpret different magnitudes as perceived annoyance. All sounds and magnitudes were presented in random order and was different for each test subject. In addition, randomization was applied to start time and the initial sound level of the test sounds to minimize the risk for bias and habituated behaviour.

The duration of the test lasted about 2 hours including two 5 minutes breaks per session and a longer 15 minutes between the two sessions.

#### 2.3.4. Compilation of results

Every individual judged 60 different stimuli 4 times, thus in total 240 stimuli. To correct for outliers, the analysed data are excluding results which deviated more than 10 dB from the individual median of the four repetitions of each sound excerpt. This elimination was performed in total 74 times of the 3264 stimuli (2.3%) for the single sound ratings and 114 times of total 4896 stimuli (2.4%) for the masked sound ratings. None of the participants had in total more than 8% corrected outliers (median 3%) and therefore ratings from all participants were included in the result. The individual geometrical averages were then calculated for every level and sound.

## 3. Results

### 3.1. Single sounds ratings

To assess the capacity of the subjects to estimate noise annoyance coherently the intra- individual correlation coefficients were calculated for averages of the first two and last repetitions for each stimulus and level (outliers omitted) and the average value over all individuals was 0.91, with a minimum value of 0.85 and maximum of 0.98.

Figure 2 show a scatterplot of the individual average ratings for the single sounds. Over each annoyance magnitude, displayed on the horizontal axis, the four columns show the produced sound level for the (order left to right) small-, medium-, large-WT and pink noise test sounds. The colour-filled markers in each column represents the inter-individual average while the empty markers are the listeners intra-individual average. When interpreting the graph, the authors wish to emphasize that the graph might be counterintuitive. Observe that; high sensitivity to the sounds appears as a low ratings of sound levels and vice versa; low sensitivity to the sounds appears as a high ratings of sound levels i.e. sounds are tolerated at higher levels before a specific annoyance is reached. As expected, the sound level increases with the increased level of annoyance and the inter-individual average is approximately linear with exception of a slight increase between the levels “weak” and “moderate” corresponding to a step from 37 to 47 dB(A). Relevant and realistic exposure levels for wind turbine noise at nearby dwellings are shown as the shaded area between 35 and 45 dB(A) and is included to highlight various wind turbine noise guidelines. As an example, the Swedish guideline recommend noise levels not to surpass 40 dB(A) (outside the proximity of a residence) (Naturvårdsverket) and the Danish guideline is 44 dB(A) [15] while the WHO-guideline recommends 45 dB Lden [23]. Note, that the WHO recommendation is expressed in Lden which cannot directly be translated to equivalent sound levels. Viewing the inter-individual average, it is observed that sound levels that are corresponding to the

guideline values were perceived as on average “weakly” annoying. It is also noted that the scatter of results show that the most sensitive 10-20% of the listeners’ rate sound levels at the guideline values to be perceived as “strong” on the annoyance scale. This observation could be compared with the WHO 10% “Highly Annoyed” guideline recommendation [23] showing that the short term noise annoyance estimates in this paper is reasonably similar to the 45 dB Lden, although more listeners would be favourable to strengthen this observation. It is also noted that the intra-individual averages diverge for higher annoyance levels. For the “just noticeable” and “very weak” levels the inter-individual standard deviation is approximately 4 dB(A), while for the “strong” and “very strong” annoyance levels, the standard deviation has increased to approximately 12 dB(A). Comparing the sounds small differences are observed between the inter-individual averages and the ranking also changes for the different annoyance levels, which is interpreted as the different wind turbine noises are assessed reasonably equal. Statistical analysis, in the form of a repeated measures ANOVA with annoyance magnitude and sound type as independent variables, reveals that the rated differences are not statistically significant ( $F(3,93) = 1.869, P > 0.140$ ). Thus, based on this sample it cannot be concluded that the sounds statistically differ from each other regarding perceived annoyance.

In addition to A-weighted sound levels, the loudness levels for all 3264 individual ratings were also calculated with the model by Glasberg and Moore [9]. In Figure 3 the inter-individual averages with corresponding standard deviations for these calculations are shown. The result exhibit a linear trend with approximately 10 phon between each annoyance level which coincides with the doubling of tonal loudness as noted in (Fastl and Zwicker 1999, Fastl 1999), which is in agreement to the design of the Borg CR100-scale where a doubling in perceived stimuli intensity should equal a step between two adjacent verbal anchors [5]. Although the dynamics of the loudness scaling is plausible this alone does not necessarily imply that loudness is an appropriate predictor of perceived annoyance. If loudness was a suitable predictor of annoyance, the sounds in Figure 3 should converge. This is also the observed trend, except for the small-WT, which at the lower magnitudes “just noticeable” and “very weak” is rated about 5 phon lower than the other sounds i.e. the small-WT is perceived as more annoying at low intensities. The rated difference is statistically significant ( $F(2.033, 65.050) = 27.691, P < 1.5 \cdot 10^{-4}$ ) which was not observed for the A-weighted sound levels in Figure 2. The authors wish to emphasize that when comparing the loudness curves in Figure 3 to the A-weighted sound levels in Figure 2 the latter show less divergence thus indicating a more suitable predictor of perceived annoyance. Moreover, comparing the dynamics of the loudness scaling (40-85 phon) to the A-weighted (20–65 dB(A)), they are observed to be similar except a 20-unit bias between the two ranges. Two suggested explanations to this discrepancy by the authors are; the first is that the dB(A)-values are averages between left

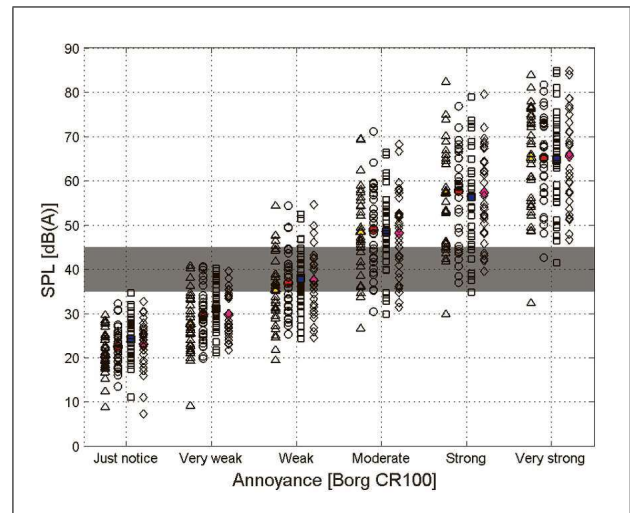


Figure 2. (Colour online) Scatter plot of intra-individual average single sound ratings for the small WT ( $\Delta$ ), medium WT ( $\circ$ ), large WT ( $\square$ ) and pink noise ( $\diamond$ ), colored markings show the average value. The abscissa shows the annoyance level and sound level (dB(A)) is shown on the ordinate.

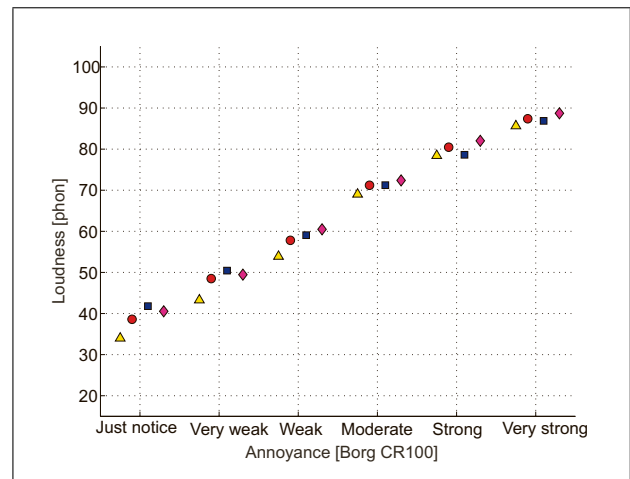


Figure 3. Intra-individual averages loudness (Phon) as a function of the produced annoyance levels of the single sound ratings for the small WT ( $\Delta$ ), medium WT ( $\circ$ ), large WT ( $\square$ ) and pink noise ( $\diamond$ ). The abscissa shows the annoyance level and sound level (dB(A)) is shown on the ordinate.

and right ear, whereas the loudness estimation is a summation between the two channels which results in a higher value (in the range of 10 phon). The second factor is that phon and dB(A) are two loudness measures that weight the sounds differently and therefore could result in different absolute values. Moreover, the size of the above discrepancy has previously been observed when comparing loudness levels to dB(A)-values [18] for traffic noise.

### 3.2. Partial masking ratings

In the second session of the test the listener was asked to produce annoyance levels for the wind turbine sounds when heard together with different background noises. This proved to be a more difficult rating task to repeat for

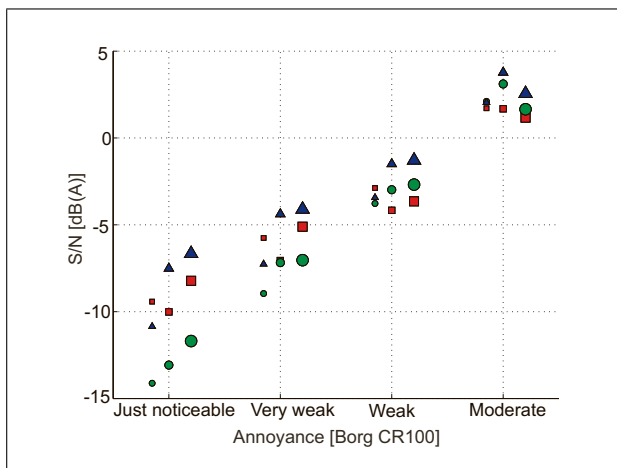


Figure 4. Shows the inter-individual average A-weighted signal-to-noise ratio as a result of produced annoyance levels for the WT sounds in the presence of the ambient background sounds. The size of the markers indicate WT size i.e. small, medium and large, while the marker shapes indicate the background sound: road ( $\square$ ), forest ( $\circ$ ) and city ( $\triangle$ ). Estimates was produced for four magnitudes from “just noticeable” to “moderate”.

the participants compared to the stimuli heard alone which was reflected in an average intra-individual correlation coefficient of 0.80 (min 0.53 and max 0.94), calculated as the average between the first and last two repetitions of each produced level.

Figure 4 shows the intra-individual average signal-to-noise (S/N)-ratios as a function of annoyance for the wind turbine sounds when heard together with the background sounds. The S/N-ratio is defined as the sound level of the wind turbine minus the sound level of the background noise. At the annoyance levels “just noticeable” the spread of the ratings is significant with ratings ranging from -7 to -14 dB(A). At the annoyance levels “weak” and “moderate” the mean S/N-ratios have decreased to -2 and 2 dB(A) and it is observed that the differences between the sounds have become much smaller i.e. when the wind turbine sounds become louder they are perceived similarly annoying, which is consistent with for example the observations of other environmental noise sources in (Berglund, Preis and Rankin 1990). For “just noticeable” annoying the mean S/N-ratio is -10 dB which is similar to the detection thresholds in [4] of -12 to -8 dB(A) and thus annoyance thresholds seem to coincide with the detection thresholds. Moreover, a reduction of perceived loudness could be observed up to +2 dB(A) S/N-ratio in [4] and agree with where the WT sounds in the present test start to be perceived as moderately annoying and consequently partial masking effects for higher annoyances is suspected to be insignificant. When comparing the different sounds in Figure 4 it is noted that sounds are clustered by WT size and background noise type, for instance, the large-WT are rated with the highest S/N-ratio for each category of background noise indicating a higher tolerance to the large-WT sound in combination with all backgrounds compared to the other WTs. The city background noise

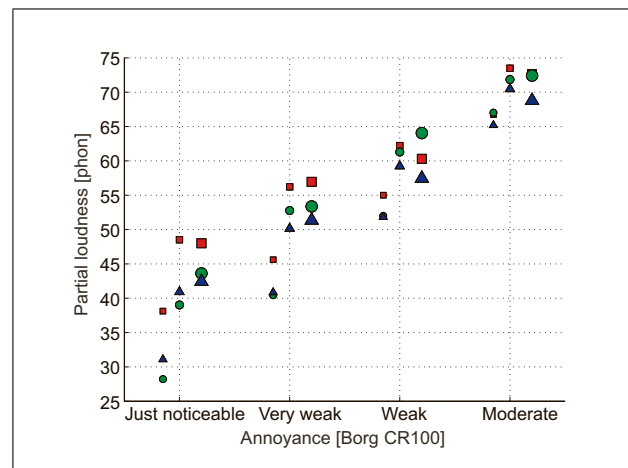


Figure 5. Shows the inter-individual average partial loudness as a result of produced annoyance levels for the WT sounds in the presence of the ambient background sounds. The size of the marker indicates WT size i.e. small, medium and large, while the marker shape indicates the background sound: road ( $\square$ ), forest ( $\circ$ ) and city ( $\triangle$ ). Estimates was produced for four magnitudes from “just noticeable” to “moderate”.

(blue markings) appears to provide the best masking effect. Investigating the differences by calculating a repeated measures ANOVA with a Greenhouse-Geisser correction applied, the differences are shown to be statistically significant ( $F(4.649, 153.423) = 14.098, P < 1.5 \cdot 10^{-4}$ ) and when performing a Post hoc test using a Bonferroni correction the significant differences is revealed to mainly be between the backgrounds. Thus, the main effect in altering the level of annoyance is contributed to the environmental background noise rather than the type of WT.

Figure 5 shows the average inter-individual partial loudness levels as estimated with the Glasberg and Moore model [9]. As observed the small-WT show continuously lower loudness thresholds indicating increased sensitivity to this sound compared to the other WTs and for lower annoyance the road background seems to improve masking compared to the other backgrounds. Step sizes of 10 phon between consecutive verbal anchors are observed between the loudness estimate, in Figure 3, and partial loudness estimates, in Figure 5, thus ensuring similar dynamics by the loudness scale. Furthermore, the average dynamic range in the two figures are also similar, ranging from 40 at “just noticeable” to 70 phon at “moderate” annoyance. The agreement in both dynamic range and threshold level for the loudness and partial loudness estimates shows that the Glasberg and Moore model is providing consistent results. However, comparing the partial loudness estimates in Figure 5 to the S/N-ratios in Figure 4 it is observed that the spread of the rated sounds is wider for the partial loudness estimates. At “just noticeable” the rating difference is 20 phon, which with the interpretation of the phon would correspond to over a quadrupling in perceived loudness. In addition, it is also observed that the order of the sounds is different between the two Figures 4 and 5. An example of this is that the city background sound

grades as the least efficient masker for estimated partial loudness while it grades as the most efficient masker if the sounds are A-weighted. Analysing the result with a repeated measures ANOVA with a Greenhouse-Geisser correction it is found that the loudness levels are statistically different ( $F(4.559, 150.444) = 118.802, P < 1.5 \times 10^{-4}$ ). Furthermore, a Post hoc analysis show that significant differences exist between the wind turbine sounds and not only between the background noise, as was the case with the S/N-ratios.

#### 4. Discussion

The perceived annoyance of the wind turbine sounds was rated using a magnitude production procedure. This rating method was chosen because it provides the listener control over the stimulus level, which by the authors is believed to have some advantages. For instance, listener fatigue is believed to be suppressed due to the active role of the listener, always being able to change the sound level and choose when assessments are produced. In this way perceived transient auditory disruptions can be avoided by letting the distractions pass, which also facilitates a less stressful test procedure as assessments are not imposed at a certain time. An indication on the validity of this method is the achieved intra-individual correlation coefficients of 0.91 and 0.8 averages for the single and masked sound ratings respectively. These numbers can be compared to a single sound wind turbine noise magnitude estimation test [2] where the average inter-individual correlation was 0.65. The above correlation coefficients indicate that magnitude production is a reliable rating procedure and that maybe in some instances facilitate estimations of annoyance for long duration single sounds and sound heard together with background noise. The main question investigated in this paper was to examine how noise annoyance from wind turbines are perceived depending on turbine size and background sound. The above presented results show inconclusive answers depending on which loudness estimate is used: A-weighted sound levels or Moore and Glasberg's time varying loudness model. Thus, to answer the main question of this paper the second objective: which model is the best predictor of annoyance, is discussed at first.

The Moore and Glasberg's loudness model for time varying sounds produced coherent results in the sense that the single and masked sound ratings, when analysed, both yielded similar ranges (40–70 phon) between annoyance levels “just noticeable” and “moderate”. When the ratings were analysed with A-weighted levels the dynamic ranges of the single and masked sound ranges differed, where the S/N-ratios were more compressed to about half the range compared to the single sounds. It can be argued that the discrepancies between the A-weighted sound levels are due to the simplified measure, where for instance temporal effects are not accounted for in contrast to the loudness model and that important information is lost. However, it is not obvious from our analysis, whether the loudness and partial loudness estimates are better predictors of annoyance when compared to the A-weighted levels and S/N-

ratios, rather it is seen that the variance of the loudness levels is higher and the A-weighted sound levels yield the most converging result, even if the higher dynamic range of loudness is considered. Thus, A-weighted sound levels appear as an adequate predictor of annoyance in the present test.

Returning to the main question and focusing on the A-weighted sound levels and in particular the masked sound ratings, it is seen that at the range 35–45 dB(A), where noise guidelines for wind turbine noise usually are stipulated and therefore pose relevant exposure levels for nearby residents, are corresponding to the inter-individual average for levels “weak” and 10% most sensitive subjects for “strong” in the present test. At these annoyance ratings the different wind turbine sounds have converged both for the single and masked sounds, indicating that there are no large differences between how annoying the different wind turbines are perceived. Thus, judging from the results presented in this paper our main question can be answered by that similar noise guidelines for different sizes of wind turbines could be recommended and that older guidelines should not have to be altered because of changing sound character of wind turbines.

Acknowledged by the authors is that differences between short-term and long-term annoyance is complex and that comparing results should be done with great care. Nevertheless, reviewing results from the two epidemiological studies [21, 12] it is seen that a steep increase in percentage highly annoyed exist between the interval 30–35 and 35–40 dB(A). Comparing to the scatter of individual ratings for the single sounds in our test it is observed that about 10% of the listeners rate the sounds at these levels as “strong” annoying. Although different scales have been used for the two types of studies, there should be similar interpretations between the words “strong” and “highly” annoyed that would imply on a possible correspondence between the produced sound levels in short-time tests and the long-time exposure levels experienced in epidemiological studies.

#### 5. Conclusion

At the sound levels corresponding to the range 35–45 dB(A), where noise guidelines for wind turbine noise usually are stipulated, no large differences between how annoying the sounds of different sized wind turbines are perceived, were found in the analysis of this listening test. The result in this paper therefore suggest that similar noise guidelines for different sizes of wind turbines could be recommended. Magnitude production was adopted as the rating procedure and results showed high correlation (above 0.8) between repeated measurements, indicating that the listeners performed ratings with reliable accuracy. In analysing the rated sounds as A-weighted sound levels and loudness and partial loudness estimates, it could not be showed that loudness is a superior predictor of short time annoyance as both the single sound and masked sound ratings exhibited a more divergent results than the A-weighted levels.

## Acknowledgements

The authors wish to acknowledge the financial support of the EU FP7 SWIP project (Grant agreement no.: 608554). The Marcus Wallenberg Laboratory are acknowledged for the recording equipment as well as for the facilities to conduct the listening test. The authors also wish to thank Prof Em Brian C.J. Moore and Brian R. Glasberg at the University of Cambridge for sharing software calculating the loudness model used to analyse this listening test.

## References

- [1] B. Berglund, A. Preis, K. Rankin: Relationship between loudness and annoyance for ten community sounds. *Environment International* **16**(4-6) (1990) 523–531.
- [2] Bolin, K., G. Bluhm, M. Nilsson: Listening Test Comparing A-Weighted and C-Weighted Sound Pressure Level as Indicator of Wind Turbine Noise Annoyance. *Acta Acustica united with Acustica* **100** (2014) 842–847.
- [3] K. Bolin, A. Kedhammar, M. E. Nilsson: The Influence of Background Sounds on Loudness and Annoyance of Wind Turbine Noise. *Acta Acustica united with Acustica* **98**(5) (2012) 741–748.
- [4] K. Bolin, M. E. Nilsson, S. Khan: The Potential of Natural Sounds to Mask Wind Turbine Noise. *Acta Acustica united with Acustica* **96**(1) (2010) 131–137.
- [5] G. Borg, E. Borg: A new generation of scaling methods: level anchored ratio scaling. *Psychologica* **28** (2001) 15–45.
- [6] C. Dickson, K. Bolin: Continuous judgment by category-ratio scaling of aircraft noise. *Applied Acoustics* **84** (Supplement C) (2014) 3–8.
- [7] ETSU-R-97: The assessment and rating of noise from wind farms. Retrieved November 19, 2019, from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/49869/ETSU\\_Full\\_copy\\_Searchable\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49869/ETSU_Full_copy_Searchable_.pdf).
- [8] H. Fastl, W. Zwicker: *Psycho-acoustics. Facts and Models*. Springer, Heidelberg, Germany, 1999.
- [9] B. R. Glasberg, B. C. J. Moore: Development and evaluation of a model for predicting the audibility of time-varying sounds in the presence of background sounds. *Journal of the Audio Engineering Society* **53** (2005) 906–918.
- [10] H. Z. Hafke-Dys *et al.*: Noise Annoyance Caused by Amplitude Modulated Sounds Resembling the Main Characteristics of Temporal Wind Turbine Noise. *Archives of Acoustics* **41**(2) (2016) 221–232.
- [11] S. Lee, K. Kim, W. Choi, S. Lee: Annoyance caused by amplitude modulation of wind turbine noise. *Noise Control Engineering Journal* **59**(1) (2011) 38–46.
- [12] D. S. Michaud, *et al.*: Exposure to wind turbine noises Perceptual responses and reported health effects. *Journal of the Acoustical Society of America* **139**(3) (2016) 1443–1454.
- [13] H. M. E. Miedema, H. Vos: Noise annoyance from stationary sources: relationships with exposure metric day-evening-night level (DENL) and their confidence intervals. *Journal of the Acoustical Society of America* **116**(1) (2004) 334–343.
- [14] H. M. Miedema, E. V. Henk: Noise annoyance from stationary sources: relationships with exposure metric day-evening-night level (DENL) and their confidence intervals. *Journal of the Acoustical Society of America* **116**(1) (2004) 9.
- [15] Miljøministeriet: Miljøministeriets bekendtgørelse nr. 1284 af 15. december 2011. Retrieved November 19, 2019, from <https://www.retsinformation.dk/Forms/R0710.aspx?id=139658>.
- [16] H. Møller, C. S. Pedersen: Low-frequency noise from large wind turbines. **129**(6) (2011) 3727–3744.
- [17] Naturvårdsverket: Riktvärden för ljud från vindkraft. Retrieved November, 2019, from <https://www.naturvardsverket.se/Stod-i-miljoarbetet/Vagledning/Buller/Buller-fran-vindkraft/buller-vindkraft-riktvarden/>.
- [18] M. E. Nilsson: A-weighted sound pressure level as an indicator of short-term loudness or annoyance of road-traffic sound. *Journal of Sound and Vibration* **302**(1) (2007) 197–207.
- [19] no.92-1444: L. Loi no.92-1444 du 31 décembre 1992 relative à la lutte contre le bruit (Act no. 92-1444 of 31 december 1992 on noise prevention).
- [20] E. Pedersen, F. van den Berg, R. Bakker, J. Bouma: Can road traffic mask sound from wind turbines? Response to wind turbine sound at different levels of road traffic sound. *Energy Policy* **38**(5) (2010) 2520–2527.
- [21] E. Pedersen, K. P. Waye: Perception and annoyance due to wind turbine noise—A dose–response relationship. *Journal of the Acoustical Society of America* **116**(6) (2004) 3460–3470.
- [22] K. P. Waye, E. Öhrström: Psycho-Acoustic Characters of Relevance for Annoyance of Wind Turbine Noise. *Journal of Sound and Vibration* **250**(1) (2002) 65–73.
- [23] WHO: WHO Environmental Noise Guidelines for the European Region. Retrieved November 19, 2019, from [http://www.euro.who.int/\\_\\_data/assets/pdf\\_file/0008/383921/noise-guidelines-eng.pdf?ua=1](http://www.euro.who.int/__data/assets/pdf_file/0008/383921/noise-guidelines-eng.pdf?ua=1).
- [24] S. von Hünnerbein, A. King, B. J. Piper, M. Cand: Wind turbine amplitude modulation: Research to improve understanding as to its cause and effect—Work Package B(2): Development of an AM dose-response relationship. from <https://www.ref.org.uk/Files/RUK-B2.pdf>.