The Swedish Environmental Protection Agency (Naturvårdsverket) recommends that the sound level from wind turbines in Sweden should not exceed 40 dBA in residential zones (or 35 dBA if low sound levels are desired). Also, if the difference between A-weighted and C-weighted sound levels exceeds 20 dB, the Swedish EPA recommends that the low frequency content of the noise is assessed and the guidelines for indoor low-frequency noise are considered. Furthermore, if the sound contains clearly audible tones, the value should be 5 dBA lower.

In 2015 the Swedish EPA funded a project, which was triggered by the trend of increasing wind turbine size and height. The project was aimed at assessing parameters for prediction models of wind turbine noise annoyance. In one of the project WP, a number of socioeconomic factors were assessed by means of a questionnaire, which was addressed to residents in the vicinities of seven wind turbine parks with relatively large wind turbines. The responses were mapped to estimations of the respective noise exposures. In another WP, recordings of wind turbine sound were made. A number of typical characteristics of the wind turbine noise were selected, and a listening test was performed to study whether the respective characteristics had any influence on the subjective annoyance, and to what extent. All listening stimuli were based on the same recording, with only one characteristic changed at a time, and the output normalized to the same A-weighted sound level. The characteristics included low-frequency content, tonality, modulation and sound level. The results are presented together with a discussion of future work needed to formulate a prediction model for wind turbine noise annoyance.

1 Introduction

1.1 Background

It is a well-known fact that community noise is an increasing environmental problem which is known to cause adverse health effects (e.g. [1]). In comparison with other sources of sound such as traffic noise, wind turbines give rise to a higher degree of annoyance given the noise level (e.g. [2]). As a consequence, the problem of environmental noise from wind turbines has been addressed in a large number of studies worldwide. In particular, several studies have been funded by the Swedish EPA (Environmental Protection Agency, in Swedish Naturvårdsverket) in order to account for Swedish conditions [3]. It has been shown that there is a dose-response relationship between A-weighted SPL and the frequency of annoyance [4], and this forms a base for noise regulations. However, in these studies no relationships to other annoyance factors have been reported for wind turbine noise, besides the dose-response relationship. Therefore, in the recommendations from the Swedish EPA and the Swedish National Board of Housing, Building and Planning (Boverket) it is first and foremost stated that the sound level from wind turbines in Sweden should not exceed 40 dBA in residential zones (or 35 dBA if low sound levels are desired). In addition to this, general practices for environmental noise are recommended. E.g., if the difference between A-weighted and C-weighted sound levels exceeds 20 dB, the Swedish EPA recommends that the low frequency content of the noise is assessed and the guidelines for indoor low-frequency noise are considered. Also, if the sound contains clearly audible tones, the value should be 5 dBA lower.
1.2 The scope of the presented project

In the last few years, there has come a new generation of wind turbines with an apparent increase of performance and a corresponding increase of height and size. The project aimed to establish whether there was a risk that the large wind turbines (> 90 m rotor diameter) would lead to a higher degree of annoyance among people living in their vicinity. The assignment included both a questionnaire and a listening and mapping of the noise levels in affected areas.

1.3 Project overview

The presented project was organised in three WP and involved six wind farms in Sweden with large wind turbines (Skaveröd-Gurseröd, Jädrås, Brattön, Malö, Möckelsjöberget and Årjäng). WP1 addressed data collection in the form of measurements according to applicable parts of the IEC6140-11 standard and recordings at the wind farms.

In WP2 the recordings were used to design a set of sound files for two listening tests. One test addressed annoyance from various characteristics of the files, and the other test assessed how different various characteristics were perceived.

In WP3 a questionnaire were sent to respondents living in the areas affected by the noise of the selected wind farms. The questions pertained to the perception of the wind turbine noise and the socio-economic situation.

2 The Listening tests

2.1 The selection of test files and test panel

From the recordings of WP1 a set of 13 files were selected as a representative span of the total set of recordings. The selected set was evaluated in a workshop where potential perceived characteristics were identified and listed. Identification process involved listening in combination with analysis of corresponding audiograms. The list was compared with what had been reported in the literature for the new generation of wind turbines, e.g. distinct amplitude modulation, “swishing” sound etcetera [8], [3], [9].

In order to create realistic sounding test files for the listening tests, all files were based on one single file, which was selected as being typical for all characteristics of the revised list. Also, all the modifications were based on the already present characteristics in the file. The selected characteristics were then identified in the selected file, isolated with filters, and extracted into separate files. Two distinct tones were selected for assessing tonality: One high-pitched narrow-band tone at 1315 Hz, and one low tone, at 156 Hz, with a much more broadband character. The 1315 Hz tone did not coincide with any of the harmonics of the 156 Hz tone (only the first two harmonics were detectable in the audiogram, anyway), and was clearly audible. The low-frequency noise was singled out with a low pass filter at 140 Hz.

In the case of the modulation, the phase differed in different frequency bands. A high-pitched (over 1 kHz) modulated sound, which may be described as “swishing”, was present in a broad band with a centre at approximately 3 kHz, and came approximately 0.5 s after a lower pitched modulated sound (predominantly sound energy below 400 Hz) centred at approximately 300 Hz, which may be described as “wooshing” (although there is no unambiguous definition in either case). The corresponding frequency bands were extracted into separate files and Hilbert transformed to obtain the amplitude envelopes. The difference in phase of the amplitude modulation for the two bands can be seen in Figure 1. The amplitude envelopes were then used to magnify the existing modulation in the respective bands, and were utilised together to enhance the overall modulation (with various phase) over a broader frequency range.
The listening files were then created by mixing the original file with the respective isolated characteristics files. A high frequency boost was suggested as a potential characteristic to assess, but was later rejected on basis of a pilot study, since it was not perceived as sounding natural. In addition to the characteristics above, also the influence of change in SPL (sound pressure level) was assessed. Today, A-weighted SPL is the most commonly used indicator of annoyance. In order to include A-weighted SPL as a parameter in the assessment, all modified files were normalised to the same A-weighted SPL, and in addition one file was created in which the A-weighted SPL of the original file was raised with 2 dB. The updated list, with the selected characteristics and corresponding listening files is presented in Table 1.

### Table 1: Characteristics selected for the listening test

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Corresponding physical property or parameter</th>
<th>Selected sound file</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>Reference - no modification</td>
<td>1) Original</td>
</tr>
<tr>
<td>Tonality</td>
<td>Audible tones with bandwidth much smaller than a 1/3 octave band</td>
<td>2) 156 Hz tone, 3) 1315 Hz tone</td>
</tr>
<tr>
<td>Low frequency</td>
<td>Noise in the frequency range below 140 Hz</td>
<td>4) Low frequency</td>
</tr>
<tr>
<td>Modulation</td>
<td>Amplitude modulation, complete and in different frequency bands, with modulation frequencies in the range 0.25 Hz &lt; f &lt; 2 Hz</td>
<td>5) Modulation @ 300 Hz, 6) Modulation @ 3000 Hz, 7) Complete modulation</td>
</tr>
<tr>
<td>SPL</td>
<td>A-weighted sound pressure level</td>
<td>8) Sound level</td>
</tr>
</tbody>
</table>

A test panel of 25 persons was selected with an appropriate spread in gender and age (from 25 to 62). The participants were performing a paired comparison test and thereafter a semantical differential test in a sequence, and the test times for both tests were typically around 15-20 minutes. The oral instructions to the participants were that the sound pertained to wind turbine noise, and that they should keep the headphones on during the whole test. The sound level was adjusted to match the recommended maximum sound pressure level in residential zones in Sweden (i.e. $L_A=40$ dBA).

At the beginning of each test, there was a set of background questions, to associate the results with metadata about age, gender, known/perceived hearing deficits and also to what extent the participants themselves had been annoyed by wind turbine noise during the last 12 months.
2.2 The paired comparison test

The first listening test involved the eight sound files combined in pairs in all possible combinations. The order of the comparisons was randomized by the software for the test design and data acquisition.

The test included a written instruction to the participants that 28 pairs of sound files would be presented to them, and in each case they should listen to both sound files presented in a random order selected by the test program, and choose which one of them was perceived as being the most annoying (it was possible to repeat listening to one or both of the files before the choice was made). They could also choose a third alternative, that both sounds were perceived as equally annoying, or that no difference could be detected between the both sounds. However, the test design also included some control checks in the form of reversal of orders and comparisons of the same listening files, which resulted in sets of incomplete comparisons. After all participants had concluded the tests, the test data was evaluated with a Bradley-Terry-Luce model, which resulted in a parameter set giving the relative annoyance of the samples in the set and an error covariance matrix from which a 95 % confidence was calculated for each of the samples, shown in Figure 1.

Figure 2 Relative annoyance with 95 % CI. The assessed characteristics were (cf. Table 1) 1) Original, 2) 156 Hz tone, 3) 1315 Hz tone, 4) Low frequency, 5) Modulation @ 300 Hz, 6) Modulation @ 3000 Hz, 7) Complete modulation and 8) Sound level

Table 2 List of the seven qualities which were evaluated in the semantic differential test, for each of the sound files

<table>
<thead>
<tr>
<th>Objective</th>
<th>Aspects</th>
<th>Number of scale steps</th>
<th>Low and high delimiters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluate the emotional effects of the sound on the test subjects</td>
<td>Arousal</td>
<td>9</td>
<td>“very relaxing” - “very stressful”</td>
</tr>
<tr>
<td></td>
<td>Valence</td>
<td>9</td>
<td>“very positive” - “very negative”</td>
</tr>
<tr>
<td></td>
<td>Restoration</td>
<td>5</td>
<td>“very good” - “very bad” possibilities for restoration</td>
</tr>
<tr>
<td>Evaluate to what degree the respective acoustical aspects can be perceived</td>
<td>Fluctuation</td>
<td>9</td>
<td>“very little” - “very much”</td>
</tr>
<tr>
<td>in the sound by the test subjects</td>
<td>Bassiness</td>
<td>9</td>
<td>“very little” - “very much”</td>
</tr>
<tr>
<td></td>
<td>Tonality</td>
<td>9</td>
<td>“very little” - “very much” of audible tones</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>9</td>
<td>“very soft” - “very loud”</td>
</tr>
</tbody>
</table>
2.3 The semantic differential test

In the second listening test, each of the eight sound files were evaluated in a semantic differential test, comprising seven emotional or acoustical aspects (or qualities/properties, listed in Table 2) that were tested to what extent they could be perceived in the sound files. The participants were asked to rate the aspects in scales with a number of discrete steps, where the extreme values were labelled according to Table 2.

The test data was analysed with a repeated measures ANOVA. The resulting significance levels show that there was a main effect of each of the aspects on a 95 % significance level, except for the two cases of Bassiness and Tonality, respectively. I.e. in these two cases there were no significant difference between any of the tested samples, but in case of the other aspects, at least one pair of samples had a significant difference in between them.

In addition to the analyses discussed above, also a PCA was performed, to evaluate the relation between the aspects. The first two components explains 84.4% of the variance, and the coordinates for the test files and the aspects are listed in Table 3.

<table>
<thead>
<tr>
<th>Test file</th>
<th>PC1 (61.41%)</th>
<th>PC2 (22.96%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>-1.18</td>
<td>0.18</td>
</tr>
<tr>
<td>156 Hz tone,</td>
<td>1.39</td>
<td>1.42</td>
</tr>
<tr>
<td>1315 Hz tone</td>
<td>-0.67</td>
<td>-0.44</td>
</tr>
<tr>
<td>Low frequency</td>
<td>-0.81</td>
<td>1.63</td>
</tr>
<tr>
<td>Modulation @ 300 Hz</td>
<td>-0.33</td>
<td>-0.72</td>
</tr>
<tr>
<td>Modulation @ 3000 Hz</td>
<td>-0.45</td>
<td>-0.69</td>
</tr>
<tr>
<td>Complete modulation</td>
<td>0.78</td>
<td>-0.98</td>
</tr>
<tr>
<td>Sound level</td>
<td>1.28</td>
<td>-0.40</td>
</tr>
</tbody>
</table>

Figure 3 The plane of the first two principal components with aspects. The ring indicates a cluster strongly coupled with PC1 and consisting of Arousal, Valence, Restoration and Loudness, and the ellipse includes Fluctuation and Tonality. Bassiness is located apart from the rest and strongly coupled with PC2.
A Tukey’s HSD (Honestly Significant Difference) test was performed on the data, in which only the “156 Hz tone” file had a significantly higher level of all the emotional aspects (Arousal (99%), Valence (99%) and Restoration (95%)), than the “Original” file. The only other file that had a significant higher level of an emotional aspect than the “Original” file was the “Sound level” file, but only for Restoration, and only at a 90% significance level.

3 The socioeconomic questionnaire

The questions were formulated in cooperation between two projects (the other one being a parallel project on propagation models etc) and was general about the participants: housing situation, living conditions, bedrooms orientation to the wind turbines, noise annoyance, visual impact of wind turbines, other perceived impact of the wind turbines, general attitudes towards wind energy, financial compensation from wind farms, demographics, health, sound sensitivity and negative affectivity. The questions have previously been used in research into wind turbines and were selected because they were found to have effects on the disturbance experience in earlier studies.

3.1 Data Processing

Data processing and all analyzes were performed using the SPSS statistics package. Disturbance from noise sources was evaluated with an ISO standardized question [10]. In the Swedish translation, the noise annoyance is measured on a five-degree scale from: "not at all disturbed" to "very much" [11]. The scale could thus be used in its five-grade original version.

3.2 Results

To investigate how prominent wind turbines as sources of noise annoyance, all listed noise sources in the questionnaire were compared with respect to proportion of annoyed participants. Figure 4 shows that wind turbines are the most common source of annoyance among the participants. 1.75 times more participants are annoyed by wind turbines (7.2%) than by road traffic noise (4.1%).

Figure 4 Proportion of participants annoyed by various noise sources. The noise source categories are from left to right: Wind turbines, road traffic, neighbours, construction work, maintenance (garbage collection etc), trains, airplanes, ventilation and miscellaneous.

The proportion of participants annoyed by wind turbine noise at the seven wind farms was also studied. The two wind farms in top had almost the same number of annoyed participants in their vicinity, 13.4% and 13.3% respectively. Then there was a gap to the next level of 5.3% and 5.2% respectively. Below there was one wind farms with 3.9% and one with 3.3% annoyed participants. For one of the wind farms, no participants reported to be annoyed.
For the individual perception of wind turbines, financial incentives have been shown to affect the degree of annoyance. Bakker [12] showed that approximately 7 percent of the residents who did not receive any financial gain from the wind turbines, were quite or much bothered by the wind power indoor noise. The corresponding figure for residents with some form of economic compensation was 0 percent. These and other socio-economic factors, such as the possibility of own ownership in the wind power plants, was planned to be investigated with regard to their impact on perceived annoyance in this project. Unfortunately, the participant number was too small for such an analysis to yield a significant result. Too few were affected by economic incentive models (70 people), and the analysis basis became much smaller when divided into respective subcategories models.

4 Discussion and conclusion

The listening test were designed to take into account various characteristics of modern wind farms, as described in the literature [8], [3], [9]. For example, low frequency noise has not been reported as being particularly annoying. Correspondingly, the “low frequency” file of the listening tests, with increased contents below 140Hz, had a moderate impact on the annoyance in comparison with the original file. However, in the semantic differential test it can be seen in that the Bassiness aspect of the “156 Hz tone” is assessed almost as high as in the “low frequency” file. I.e. if the low frequency content is claimed to be problematic in field studies, it may be due to tones at least as high as the 160 Hz third-octave band.

As can be seen in the PCA plot, Figure 3, the emotional aspects (Arousal, Valence and Restoration) are well correlated with the first PCA component, and are all clustered together with the acoustical aspect Loudness (as perceived in the test). It should be noted that the test file with a prominent 156 Hz tone had a higher “Loudness” rating than the test file which was 2 dB louder (the “sound level” file). I.e. even though the aspect “Loudness” had a negative impact on the emotional aspects in this study, it is a consequence of the relative loudness of the 156 Hz tone and not of the A-weighted level, which was confirmed in the Tukey’s HSD test.

The paired comparison test showed that an increase of the A-weighted sound level was not the only way to increase the annoyance in comparison to the original sound file. Keeping the A-weighted sound level constant while increasing the relative level of a low tone at 156 Hz, or the relative level of modulated noise in the range 200 - 400 Hz, had almost the same effect. But such characteristics will be more or less masked if A-weighted measurement results used at in an evaluation. This, together with the results from the semantic differential test indicate that care must be taken when using A-weighted sound levels, in particularl y in connection with self-reported annoyance of wind turbine noise.

Finally, it should be noted that the findings and implications presented here are made from a pure auditory perspective. A wide range of other factors will have a great impact, and subordinate several of the observations. As an example, the presented project was also supposed to include an analysis of socioeconomic factors based on a questionnaire. However, the socioeconomic investigation did unfortunately not contribute to the pending prediction model, since the participant basis was too small, or rather, since too few of the participants were affected by any economic compensation models.
5 Acknowledgement

This research project was funded by the Vindval Program.
The author gratefully acknowledges Penny Bergman, Fredrik Öberg, Jesper Alvarsson and Niklas Rosholm for assistance with data collection.

References


[7] Renewable UK. The Development of a Penalty Scheme for Amplitude Modulated Wind Farm Noise - Description and Justification, Report 2013


