HAMBURG 2016

A quantitative and qualitative review of amplitude modulation noise from wind energy development.

Sarah LARGE[1]

[1] MAS Environmental

ABSTRACT

The adverse impact of amplitude modulation (AM) has been acknowledged in research, papers, and anecdotal evidence since at least 2002. Only in 2013, in the UK, did wind industry acousticians finally acknowledge the common impact of AM and the need for control. Some still deny the need for control at public inquiries. Despite two independent bodies, the Institute of Acoustics (IoA) and the Independent Noise Working Group (INWG), announcing a review of AM from a planning perspective in 2014 and the UK Government's DECC commissioning further review of AM in 2015 the UK is still, at the time of writing, without any unanimous and/or accepted guidance for assessing and conditioning AM at the planning stage. This paper reviews a range of metrics and methods that have been proposed and their ability to work with real world data. This paper questions whether a "one size fits all", purely quantitative approach reflects subjective AM impact and provides initial findings of a preliminary study testing whether different manifestations of AM can be considered equal.

1. INTRODUCTION

The rhythmic variation of wind turbine noise, occurring at blade pass frequency, is generally termed amplitude modulation (AM). This aspect of wind turbine noise is commonly generated by wind turbines (1). There is still no consensus amongst those researching and assessing the occurrence of AM in the far field as to how to describe (qualitatively and quantitatively) this aspect of the noise, to what degree it is a problem, how often it occurs, what causes it and on a more basic level what term should be used to describe it. In the UK AM that occurs in the far field and is outwith the definition provided in ETSU-R-97 has been referred to as "excess" or "enhanced" amplitude modulation (EAM). Others have referred to it as "greater than expected amplitude modulation" (GTEAM) whereas a research project published in 2013 provided definitions for "normal" AM (NAM) and "other" AM (OAM) (2). This lack of clarity in the acoustics community is equally reflected by those who experience AM in and around their homes and provide a range of descriptors and comparators when describing impact (3).

Whilst wind farm noise and specifically AM are a significant source of complaints, relying on complaints alone cannot fully describe the extent of the problem (4). The WHO estimate that only 15-25% of people identified as highly annoyed by noise will complain (5). The character of AM within wind farm noise adds to its annoyance (6). In the UK most of those working for and with the wind industry appear to have recognised the issue of AM and accept it is in need of control, but how is still a matter of debate.

2. METHODS FOR MEASURING AND ASSESSING AM

2.1 Noise conditions

A minority of wind farms in the UK have been approved with a planning condition to control AM. In 2009 the Den Brook Wind Farm was approved with a condition that considered the regular occurrence of AM in excess of 3dB (as a peak to trough value in the A weighted time series) in the far field as unreasonable. This has since been added to by the developer to include a scheme for the identification of GTEAM (Den Brook Wind Farm, "condition 21") and involves finding the energy within a critical band centred on the blade passing frequency, looking for values greater than 2.5.

A draft planning condition was appended to the Renewable UK research published in December

[1] sarah@masenv.co.uk

2013 (2) and this has been adopted in some cases. This uses a similar methodology to that of the condition 21 referenced above. It finds the energy at blade passing frequency and uses this as the AM value for that period. These values are averaged over a 10 minute period followed by a further averaging of the 10 minute values for each wind speed (using a best fit line). A penalty (maximum of 5dB) is applied to the wind farm noise level depending on the AM value for each wind speed. Significant flaws with this approach and the developer modified Den Brook condition (condition 21) have been identified (7) through testing with an extensive database of EAM obtained from the Cotton Farm Wind Farm where a permanent monitoring station has been established.

Outside of the UK, draft planning guidelines for wind farms in New South Wales set a variation of 4dB(A) at the blade passing frequency as excessive amplitude modulation. This results in a 5dB(A) penalty applied to the measured wind farm noise level. Previous research by the author has shown that application of a 5dB penalty to the UK wind farm noise limits would fail to prevent adverse impact in the majority of cases tested (8).

2.2 Methods of identifying AM

Whilst there are relatively few criteria for judging the acceptability of AM there are a number of methods for identifying and quantifying AM. These have been categorised in to time domain, frequency domain and hybrid methods (9). Examples include the (original) Den Brook condition, which identifies AM by interrogating the temporal A weighted noise trace. The D_{AM} index is another time series based method described by Fukushima et al (10). The most commonly reported and tested method referenced in papers analysing AM uses fast Fourier transform (FFT) to find the energy at blade passing frequency. Other methods have included reconstructing the time series of the AM based on FFT analysis (11). Different methods for identifying AM are well described elsewhere and have not been detailed further in this paper (7,9).

2.3 Modifiers of AM perception

In the UK there appears to be a strong industry preference for an automated AM assessment method that outputs a single numerical value of AM for a defined time period. This value is based either on the energy found at blade passing frequency or with an estimate of the typical difference between the peak and trough of the AM. Annoyance from wind farm noise has also been related to the overall A weighted sound pressure level (L_{Aeq}) (2).

No method presented to date takes account of cumulative factors such as spectral content, impulsivity (though it is noted that others have investigated use of impulsivity for quantifying AM(12)) or other features that may contribute to the perception of AM. Pedersen and Waye identified descriptions of swishing, whistling, pulsating / throbbing and resounding as the most common sources of annoyance (13). Annoyance has also been identified as a function of modulation depth, modulation rise time, modulation frequency (including harmonics), the average level of the noise and psychological responses (9). The unpredictability of AM and better audibility at night, due to lower masking noise and greater transmission, have also been highlighted as factors contributing to annoyance (14).

Thus, whilst a multitude of factors have been identified as influencing AM perception, assessment methods tend to reduce this to interrogation of a single feature of AM, relating annoyance to a numerical value based only on this one feature of the AM.

3. TESTING RESPONSES TO AM

3.1 Individual differences

It is evident from personal music preferences that there is great diversity in listener's perception and discrimination of sound versus noise. Research has found that wind farm noise is more annoying at lower levels than other sources of environmental noise (12, 15). Although certain features of wind farm noise, including AM, have been highlighted as causing annoyance there is a lack of research further identifying specific character features of AM that cause an increase in annoyance. Research by the University of Salford found that annoyance rating of AM was not significantly affected by frequency content or shape of the AM waveform but was significantly affected by the frequency rate of the modulation, the overall loudness of the test sound and modulation depth (2). Despite these initial findings the overall conclusions of the research noted that annoyance was mainly related to overall level (L_{Aeq}).

Whilst the findings of the University of Salford research indicate that a single numerical value for AM, potentially related simply to the overall average decibel level, could sufficiently describe annoyance there still remains uncertainty, for example from the factors highlighted by research discussed in Section 2.3 above. The Salford research used an artificially generated AM signal which may not include many of the characteristics, including random fluctuations, found in real world AM. The use of a single numerical descriptor for assessing noise, whilst convenient, has long been acknowledged as limited (16). There is uncertainty as to how a value for AM is to be derived and whether a judgement of acceptability based on this value would adequately address the annoying features associated with AM as described in other research and in lay evidence.

3.2 How do people hear wind turbine AM?

It could be argued that approaching wind turbine noise assessment pre-armed with a knowledge of psychoacoustics and wind turbine noise manifestation is subject to experimental bias. An acoustician is familiar with formal definitions of 'hum, 'whine', 'low frequency', 'modulating', 'irregular' and whilst amongst the acoustics community these definitions hold some clarity, to the lay person the same sound could be described in a manner of different ways.

When attempting to define an assessment tool for AM, and indeed setting a level of acceptability, a number of elementary questions require consideration.

→ Does all AM sound or impact the same?

Can listeners actually perceive a difference in AM samples? For example, is there a noticeable difference between AM that is low frequency or mid frequency dominated? Is there a noticeable difference between AM that modulates by 7-8dB or by 3-4dB. Can a listener discern whether a short extract of AM, used comparatively, has intermittent AM or constant AM and even if they can, does this alter their overall judgement of the sound over a time period? Is it simply categorised at a more basic level of 'noisy' or 'not noisy'?

→ Is there a consensus amongst the population as to what is "good" and "bad" sounding AM?

If AM doesn't sound the same and there are noticeable differences, do listeners generally agree as to what can be considered positive and negative attributes of AM? A repetitive rhythm to one listener may be soothing but to another may be considered relentless and intolerable.

→ Do existing AM metrics reflect different characteristics of AM?

If listeners do perceive differences in AM samples, do metrics for AM reflect these judgements? Does a single number rating for AM adequately, albeit potentially indirectly[2], penalize samples that are rated as most intrusive and permit samples that are considered acceptable?

3.3 MAS Environmental Wind Farm Noise Annoyance Study

In the absence of any substantial research asking or investigating the above questions it is of some concern that UK guidance on rating and assessing wind turbine AM could be imminently approved. As a result, and with limited funds and time, a basic online study was developed to begin investigating these questions. The survey is accessible online at: http://www.masenv.co.uk/survey/.

Six samples considered to have varying attributes of AM were taken from data measured by the external permanent monitoring station at Cotton Farm Wind Farm, Cambridgeshire UK.[3] All samples had the same L_{A90}, a narrow range of L_{Aeq}'s and lasted for approximately 50 seconds.[4] As the main purpose of the study was to investigate discrimination of character features and annoyance between samples of AM, no reference non-modulating wind farm noise was included in the study. Only the audio of the sample was presented on the website to prevent any visual clues for annoyance rating. The

---

[2] An AM control might simply state that all AM above a modulation depth of 3dB(A) is unacceptable. If adverse character and impact do not occur when there is AM with a modulation depth below 3dB(A) then the control prevents adverse impact from noise character without specifically identifying these features or focusing on preventing their occurrence.

[3] For more information and for real time information logging noise levels from the wind farm visit: http://www.masenv.co.uk/~remote_data/. The noise monitor is located in a free field location at the boundary of residential property just under 650m from the nearest turbine.

[4] The LA90 metric is used to measure wind farm noise in the UK and is therefore used as the overall decibel level of the wind farm noise assessed. In this study all samples had an LA90 of 39dB.

samples were anonymised and the order randomised each time the page was visited.[5] Participants were asked to listen to the six clips and rate them on a scale of 0-10 for perceived annoyance. Participants were then asked a few brief questions to gain some basic information including their age and the character of the area they lived in. In order to maximise participation the study was designed to be quick and easy to complete, fairly basic and anonymous. At the end of the survey participants could leave any additional feedback or information on the study / study samples in a comments box.

3.4 Limitations

As with any online study there are limitations and factors that cannot be controlled. One of the main limitations is a lack of control over the playback used by participants to listen to samples. This would vary across headphone and laptop / computer speakers both in terms of frequency output and volume. Some responses indicated that the volume was not loud enough as samples could not be heard and other comments indicated that samples were played at too high a volume. Whilst this will have some effect on the study results it is not considered to undermine the main purpose of this study, which was to investigate whether differences between AM samples could be heard and whether samples would be perceived as more or less annoying *than each other*. Assuming that participants did not change the volume between samples, these comparisons could still be fairly made. The study format also lacked control over the range of participants completing the study. Whilst a range of ages completed the study there was a lack of participants under 30 and few living in urban areas completing the study. Whilst the lack of control over participants could be considered a limitation, the number of responses that the online survey allowed is an advantage of this approach.

3.5 AM samples

The six samples used in the study are shown graphically below. These have been labeled A-F for ease of reference in this paper and should not be confused with the randomised letters A-F that appear on the study website. Plotted on to the graphs is the 100ms L_{Aeq} (black trace, read off left hand x-axis) and the AM values for each 10 second period derived using the UK Institute of Acoustics (IoA) Amplitude Modulation Working Group (AMWG) Consultation Software.[6] Also labeled on the graphs is an approximate peak to trough value of the A weighted time series trace. This allows comparison of four different methods for quantifying AM. Method 1 (pale blue trace, read values off right hand y-axis) is a time series method for quantifying AM based on the methods used in Fukushima et al (9,10). Method 2 follows an FFT approach for quantifying AM using the frequency domain (pale green trace, read values off right hand y-axis) (9). Method 3 is a hybrid approach using a reconstructed time series founded on an FFT analysis of the original data (lilac trace, read values off right hand y-axis) (9). The 10 minute AM value for each sample period, Methods 1-3, is given in boxes in the top left hand corner of the graph. This 10 minute value is the LA10 of all the individual AM values calculated within a 10 minute period (9). The individual peak to trough values can be used for comparison against the criterion set in the original Den Brook condition. For ease of comparison all graphs use the same axis scaling. A summary of the study samples' period L_{A90} and L_{Aeq} is given in Table 1 below.

Table 1: Summary of study sample's period L_{A90} and L_{Aeq}

	Sample A 31/12/2013	**Sample B 22/06/2014**	**Sample C 08/10/2014**	**Sample D 17/05/2015**	**Sample E 02/08/2015**	**Sample F 02/12/2015**
L_{A90} (dB)	39	39	39	39	39	39
L_{Aeq} (dB)	42	44	42	42	42	41

[5] The order of the samples always appears A-F but the sample attributed as A-F changed each time.
[6] For more information see: http://www.ioa.org.uk/publications/wind-turbine-noise and particularly the AMWG Discussion Document available online:
http://www.ioa.org.uk/sites/default/files/AMWG%20Discussion%20Document.pdf

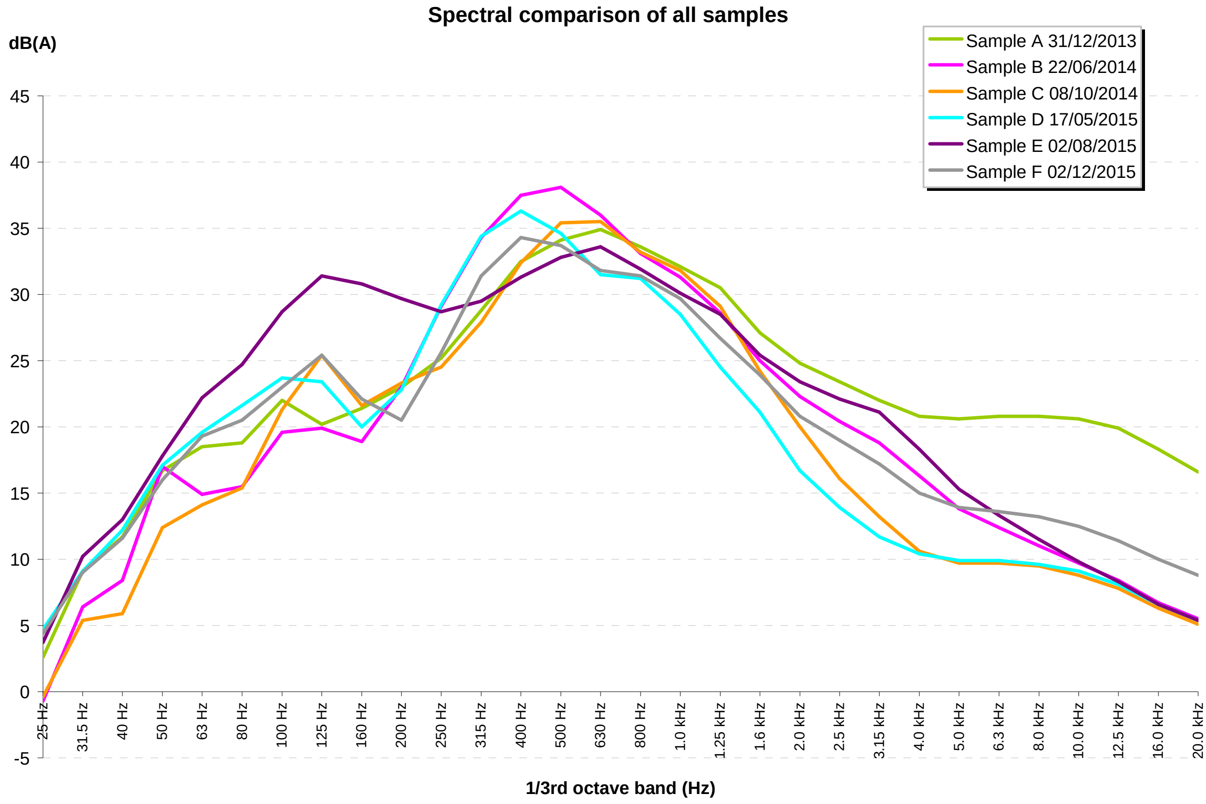

Figure 1: Spectral comparison (A weighted) of study samples.

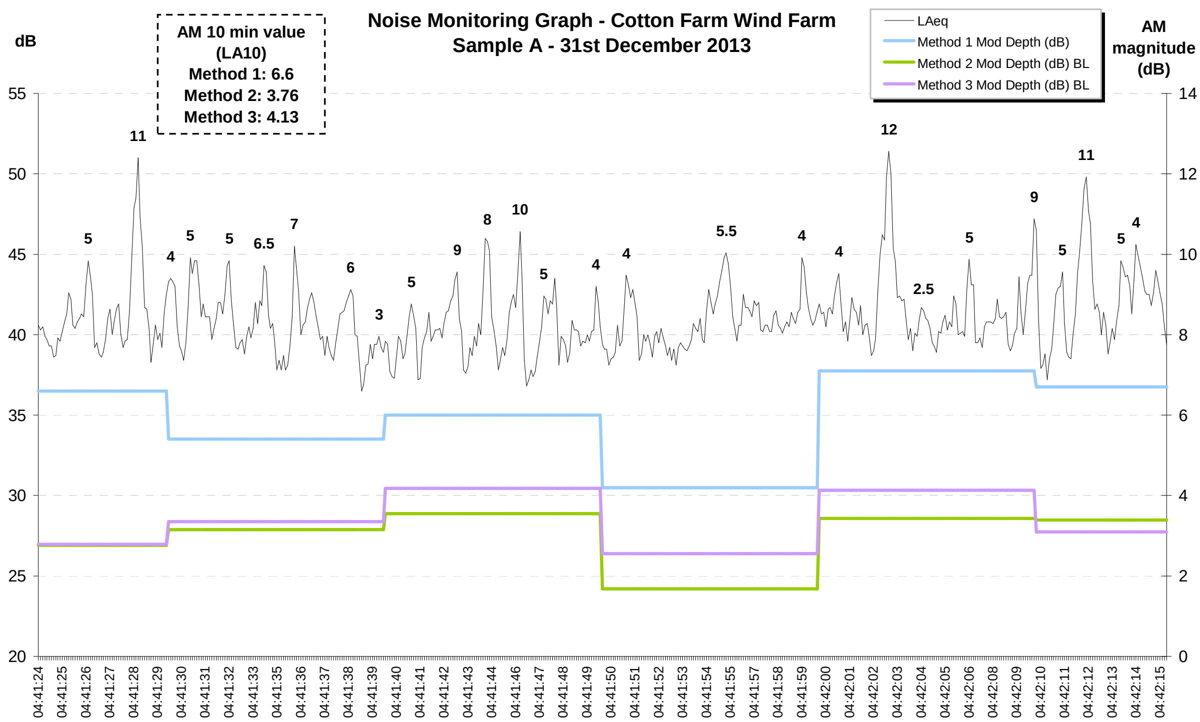

Figure 2: Study sample A - data measured on 31/12/2013. The sample on this occasion was chosen as it exhibited intermittent AM with a variable peak to trough.

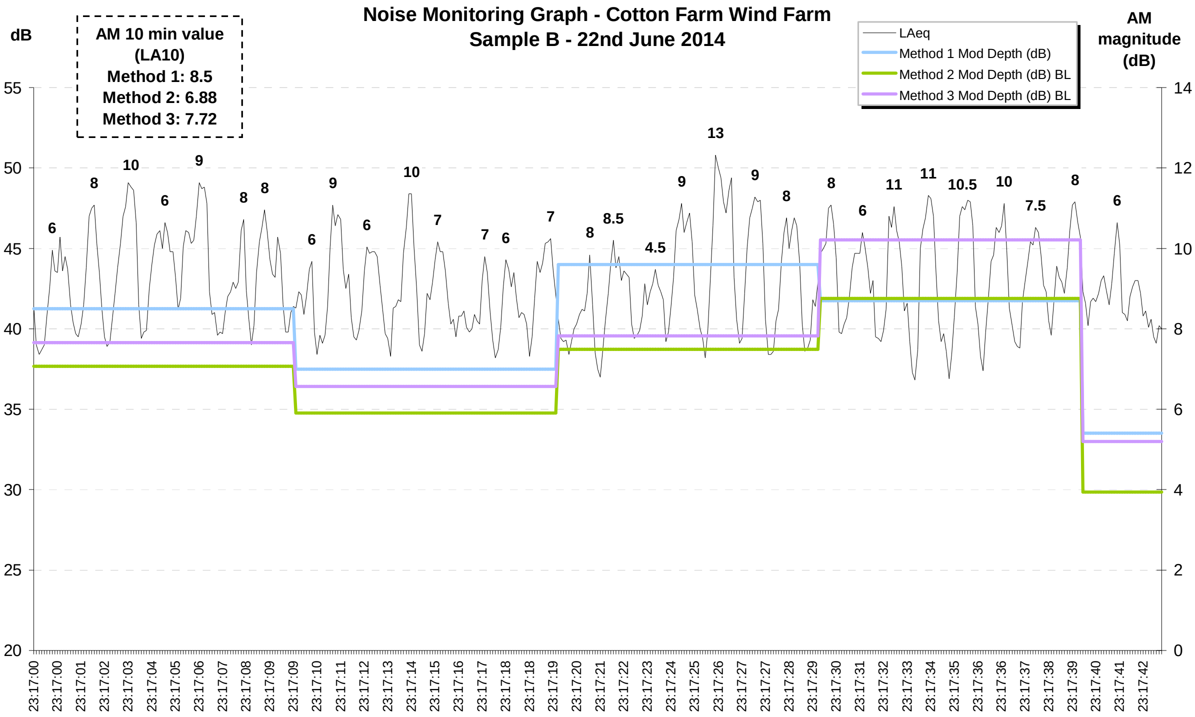

Figure 3: Study sample B - data measured on 22/06/2014. This sample was chosen as the AM was fairly consistent and constant throughout the period whilst also having a significant peak to trough variation.

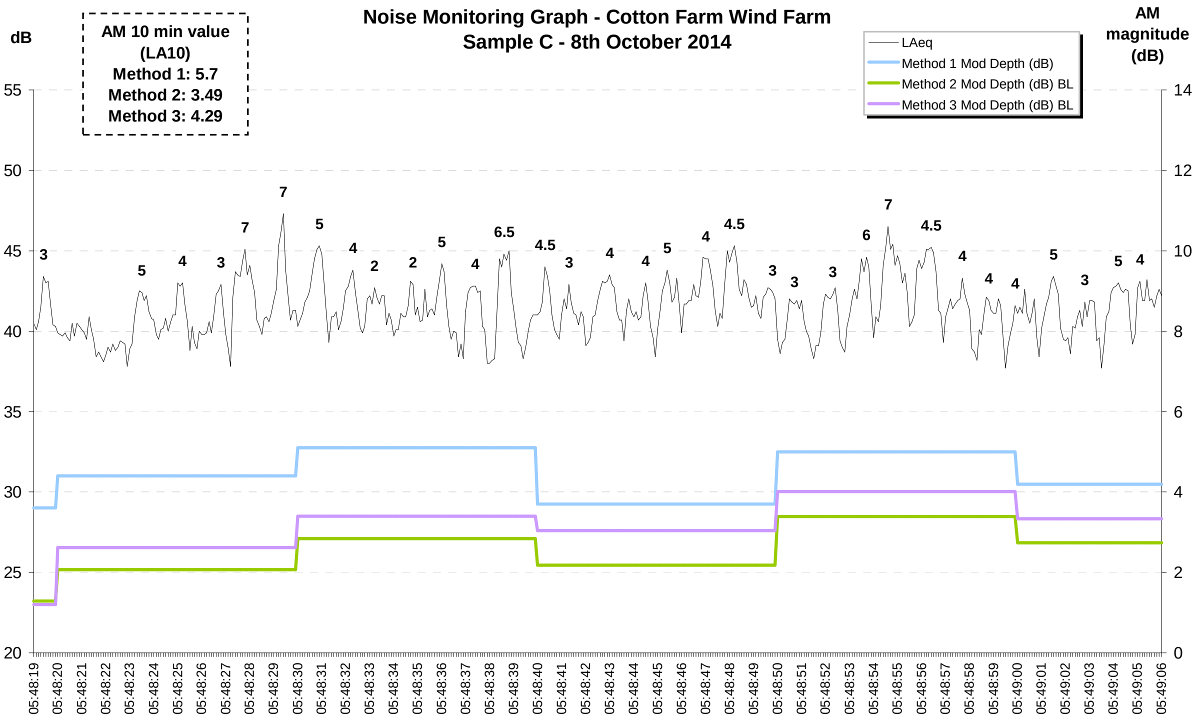

Figure 4: Study sample C - date measured on 08/10/2014. This sample is considered to show a period of AM that is not constant or consistent, with sounds of mid - high frequency i.e. "swish" rather than "whoomph" and has variable peak to trough differences.

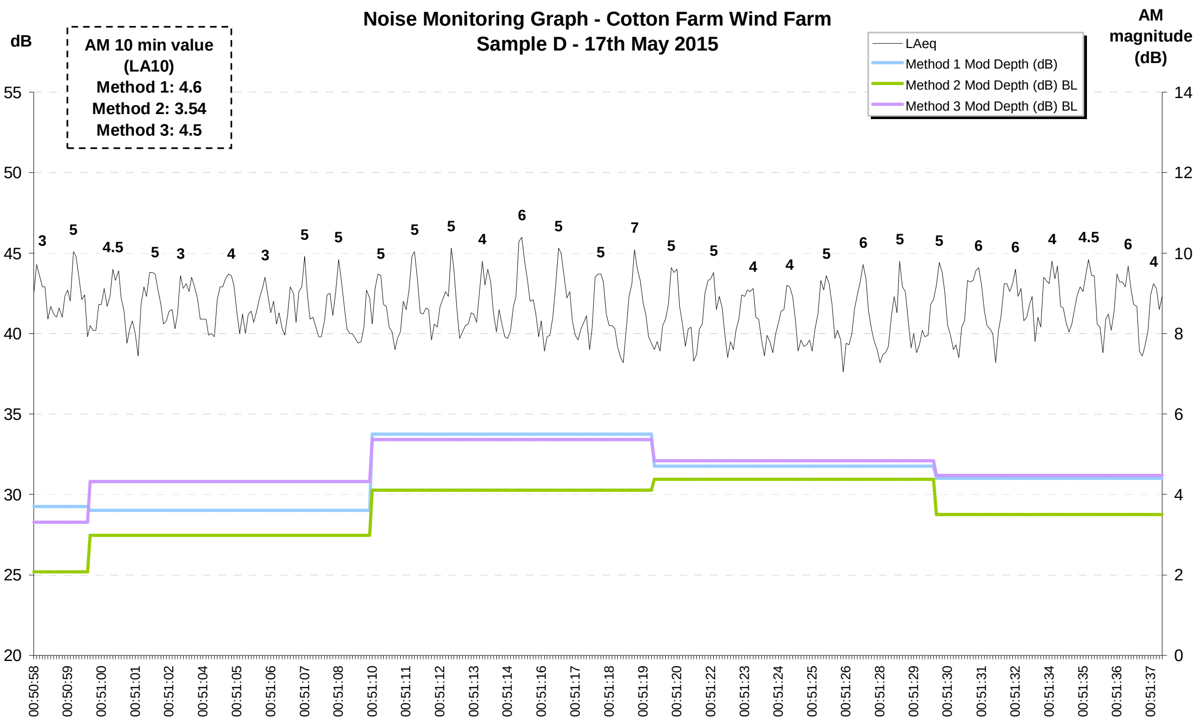

Figure 5: Study sample D - data measured on 17/05/2015. The sample on 17/05/2015 shows a period of consistent, regular and constant AM but with a lower typical peak to trough difference to that shown in the data on 22/06/2014.

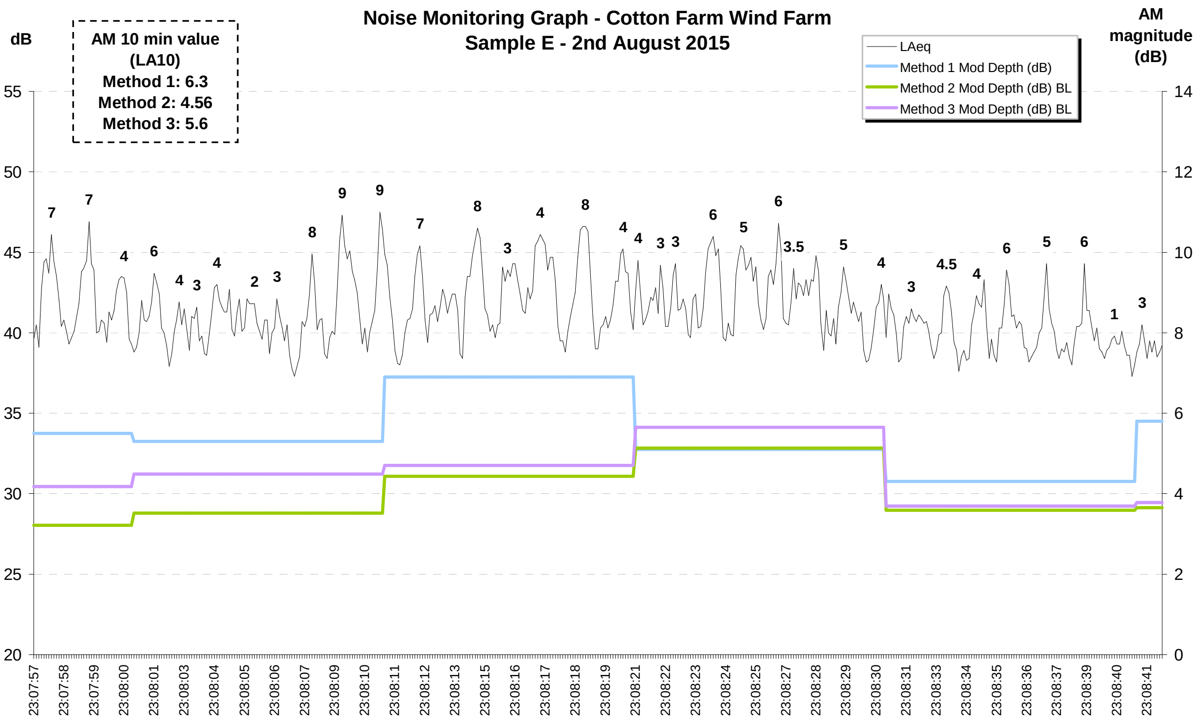

Figure 6: Study sample E - data measured on 02/08/2015. This sample was used as the AM had an intermittent and variable peak to trough difference, did not always vary significantly but had a low frequency quality to the sound, e.g. "whoomph" rather than "swish".

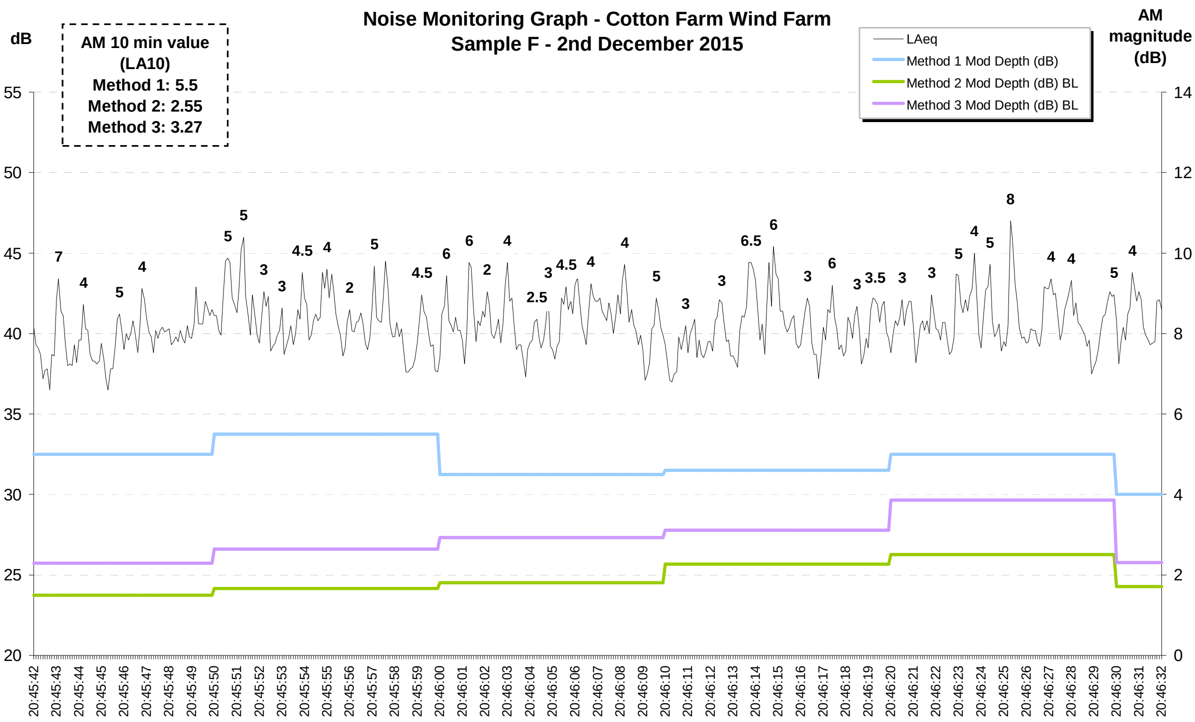

Figure 7: Study sample F - data measured on 02/12/2015. This sample was chosen as a period of lower peak to trough variation AM and also that had a mid - higher frequency sound character, e.g. "swish".

The study samples exhibit a range of differences in the modulation depth, how modulation depth depth varies with time, the character, frequency content and clarity of individual peaks and troughs and how these features vary with time. The samples represent many common features of AM including complex changes in noise character.

3.5.1 Preliminary results

At the time of writing the study had received 336 responses. A summary of the participant data is provided in Table 2 below. Preliminary analysis of the results is presented in Figures 8 and 9 below. Figure 8 shows the average annoyance ratings for all participants across each sample (see left hand axis for annoyance rating). Also plotted on the graph is the AM value derived by Methods 1-3 discussed in Section 3.5 above (see right hand axis for AM magnitude). Figure 9 shows the percentage of responses for annoyance ratings provided by participants between 0 and 10 for each sample. A line of best fit has been added for each sample to show an approximate trend in how each sample was rated. A number of participants provided additional feedback on the AM samples in the comments section at the end of the survey, including what the samples sounded like and / or why some samples were more annoying than others. Common descriptors in these responses were identified and are listed in table 3 below. A tally is provided of the number of times these descriptors were used.

No results from the initial 336 responses have been excluded from analysis. Initial analysis has not revealed any obvious outliers though further detailed analysis of individual results has not yet taken place. Full description of the study methodology and limiting factors are not discussed in this paper and will be considered in more detail post completion of full analysis.

Table 2: Summary of participant information

Age	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
Count	0	13	23	50	93	111	45	1	0

Area type	Rural	Very quiet suburban	Suburban	Residential urban	Urban near some industry	Area of heavy industry
Count	251	27	25	24	8	1

Area description	Free from intrusive noise	Impacted by some annoying noise sources on occasions	Generally noisy from transport sources	Generally noisy due to a mix of different types of noise such as industrial noise	Very noisy
Count	203	92	23	13	5

Live near wind farm?	No	Yes
Count	251	85

If "yes" to live near wind farm, do you hear wind farm noise?	Regularly	Sometimes	Never
Count	55	17	9

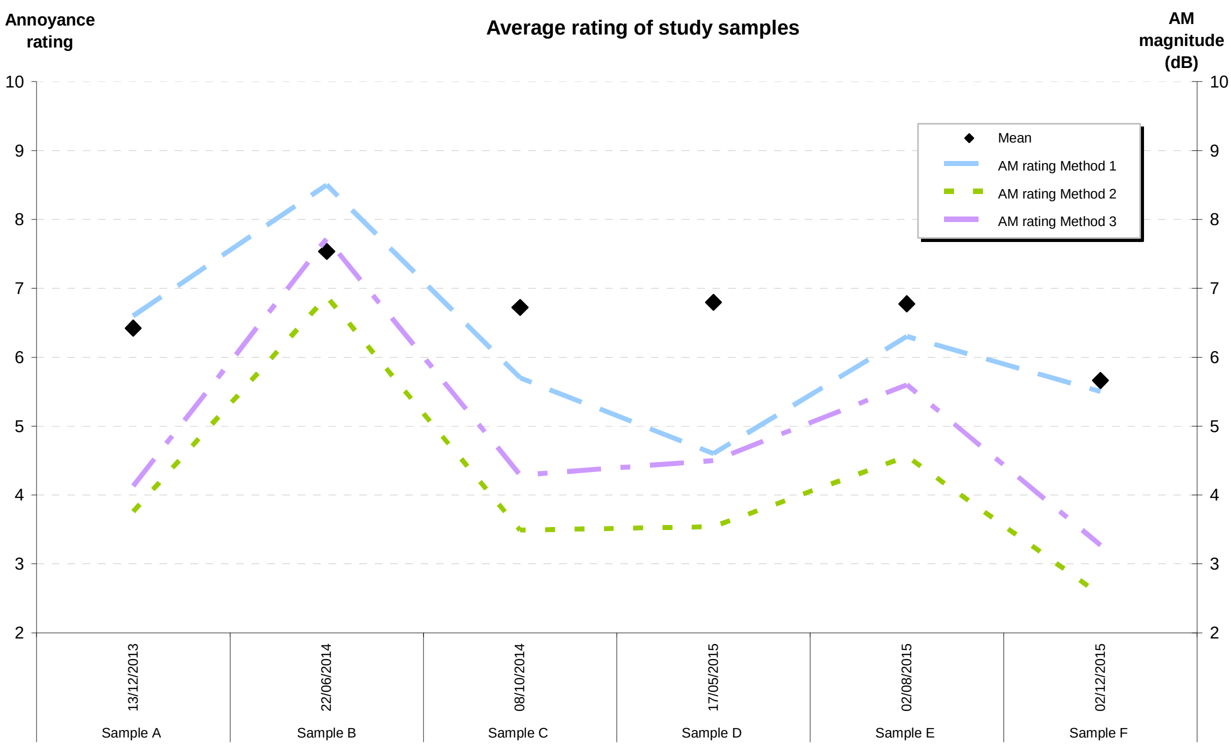

Figure 8: Average annoyance rating of study samples compared with average 10 minute AM rating for Methods 1-3.

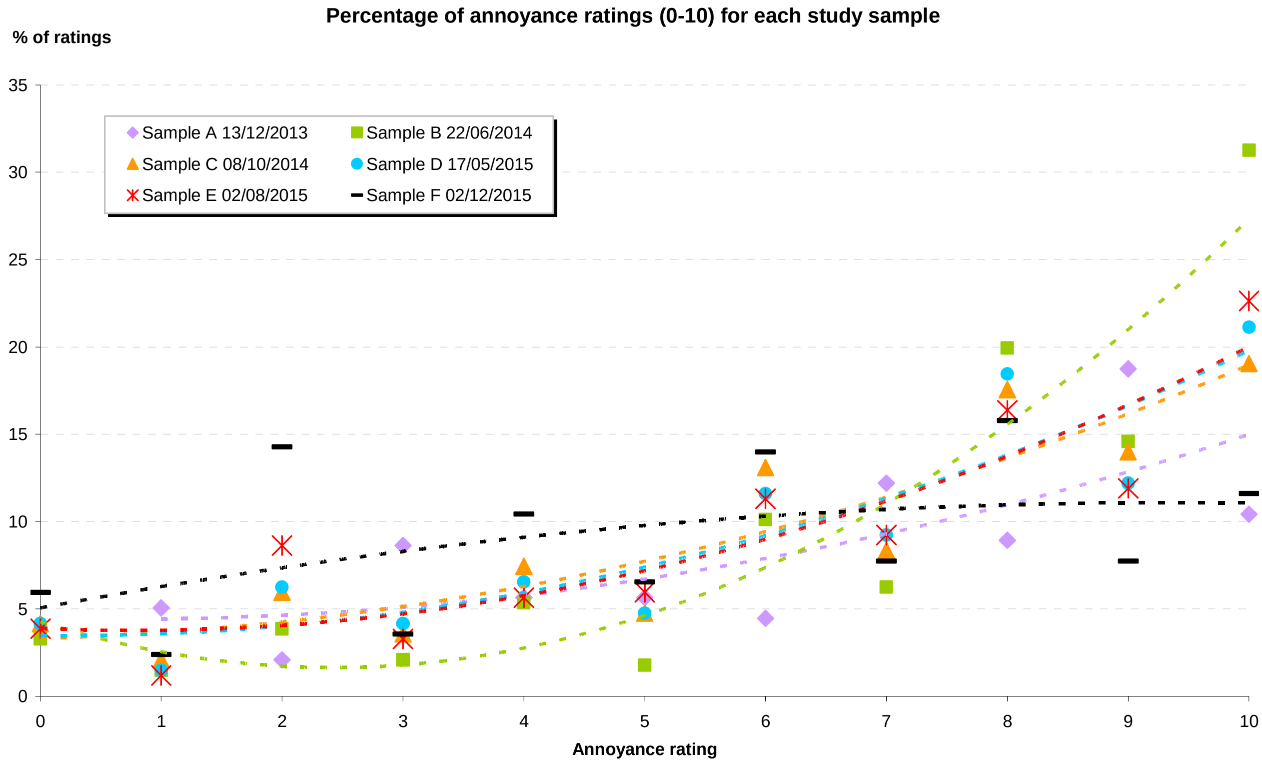

Figure 9: Percentage of responses for each study sample rated 0-10 and corresponding best fit line.

Table 3: Summary of comments relating to character descriptions of study samples.

Character feature	More annoying	Less annoying
Annoying because of the change in the noise, anticipating the noise, irregular	~~IIII~~ ~~IIII~~ III	
Continuous, regular, constant noise	~~IIII~~ I	~~IIII~~ III
Pulsing / throbbing / modulating	~~IIII~~ ~~IIII~~ III	
Repetitive noise	~~IIII~~ ~~IIII~~ II	
Rhythmic noise	~~IIII~~ III	I
Low frequency noise	~~IIII~~ II	
High frequency noise	II	
Sharper	II	
Louder	~~IIII~~ III	
Other annoying characteristics / comparisons		
Drone / hum	~~IIII~~ IIII	
Aircraft	III	
Tumble dryer	~~IIII~~	
Wind in trees	II	

4. DISCUSSION

4.1 Did participants hear differences in the samples?

The primary aim of the study was to ascertain if participants judged samples of AM with the same decibel level, as determined in accordance with UK guidance, as equally annoying.[7] The majority of participants rated the samples differently, only 46/336 participants (14%) rated the samples equally. This indicates that listeners do respond differently to samples of AM that have different characteristics and that some types of AM are more or less annoying than others.

The range of annoyance ratings given by participants was fairly small, on average ratings ranged by 3. This suggests that whilst participants judged the samples to be more or less annoying than each other, they were not judged to be significantly better or worse. This is reflected in the average scores given for each sample.

The highest annoyance rating was given to Sample B. This sample has the highest L_{Aeq} (by 2-3dB) of all the samples and has the most consistent, highest peak to trough variation of all samples. Samples D, E and C were on average all rated similarly at 6.8, 6.8 and 6.7 respectively. These samples all have the same L_{Aeq} and L_{A90} but have differing characteristics. Of particular note, Sample E has significantly more low frequency content than all other samples. Sample D consistently varies by approximately 5dB peak to trough whereas Sample E is more intermittent with peak to trough values ranging unpredictably from 2-3dB to 7-9dB. Only rated as slightly less annoying, with an average of 6.4, was Sample A. This sample is similar to Sample E but is more intermittent and has more extreme peak to trough variation, ranging from AM with 2-3dB up to 11-12dB peak to trough variation. The lowest rating of 5.7 was given to the AM in Sample F. This has a slightly lower L_{Aeq} (41dB) than the other samples and has intermittent AM with typically lower peak to trough variation. These results suggest that participants generally considered AM with consistently lower peak to trough variation to be least intrusive and AM with consistently higher peak to trough variation to be the most intrusive. However, samples that are intermittent and have a greater variation in peak to trough levels were considered to be similarly annoying compared to samples with lower peak to trough levels but which are constant and consistent throughout the period.

Figure 9 provides more information regarding how each sample was rated. It again confirms that Sample B was clearly rated as the most annoying, receiving significantly more annoyance ratings of 10 than the other samples. All samples with the exception of Sample F display a trend of increasing annoyance ratings, i.e. relatively few ratings of 0-3 and an increasing number of ratings from 4 up to 10. Sample F had a similar number of responses rating it as ‘2 – noticeable’ as it did ‘6 – annoying’. This sample does not have significant peak to trough variation compared to other samples and does not have any specific characteristics such as low frequency content or unpredictable high AM peaks (for example up to 9-11dB peak to trough variation). The lack of additional character, other than the AM alone, could perhaps explain the different overall response to this sample and more uncertainty amongst participants as to how it should be rated.

Whilst the average rating of Samples C, D, and E was more or less the same, Figure 9 does show some differences in how the samples were rated. Sample E received marginally more annoyance ratings of ‘10’ than Samples D and C. This could be taken as support that AM with more low frequency noise (Sample E) is more annoying than samples with intermittent but much higher peak to trough variation (Samples D and C). Notwithstanding that Samples B, C, D and E have significantly more ratings of '10', Sample A has a significant number of ratings of ‘9’. Sample A has the most intermittent and unpredictable AM with significant peak to trough variation of up to 11-12dB. Thus, whilst Sample A is ranked fourth out of six samples for annoyance, it still has attributes that cause participants to rate it highly for annoyance.

4.2 AM rating methods

Figure 8 plots the average annoyance rating produced by participants of the survey for each AM sample and the AM rating for each sample using Methods 1-3 discussed in Section 3.5 above and based on the 10 minute period from which the sample was taken.[8] The AM rating methods give an ‘AM magnitude’ in decibels and although it is not clear whether this magnitude is expected to reflect the typical peak to trough variation of the AM or if the value is to be used as a penalty, it is reasonable to

[7] Annoying in this discussion refers to a lay definition, in a broad sense of noisiness / unacceptability.
[8] The AM in the 10 minute periods from which samples were taken continued in a similar manner.

assume that the values may be used in this way, i.e. to rate modulation depth. As such the AM magnitude has been compared to the annoyance ratings of participants. The AM rating methods generally followed the shape of the annoyance ratings and clearly identified Sample B as the most annoying giving it an AM rating of between 7 and 9. These values tend to reflect the peak to trough values within that sample period.

Whilst the AM rating methods appear, to some extent, to reflect subjective response, the range of the 10 minute AM values attributed to the samples by Methods 1-3 is much greater than the range of annoyance ratings attributed by participants. The range of annoyance ratings given by participants across the samples is 5.7 – 7.5, indicating that whilst samples were perceived differently they were not significantly better or worse than each other. The range of 10 minute AM values for Methods 1, 2 and 3 across the samples is 4.6-8.5, 2.6 – 6.9, and 3.3 – 7.7 respectively, making some samples of AM appear significantly better or worse than others. This wider fluctuation in the empirical methods (Methods 1-3) indicates that they do not fully reflect aspects of AM that participants judge as relevant contributors to annoyance.

The lowest AM ratings were derived by Methods 2 and 3 and were given to Sample F and A. With reference to Figures 2-6, these are the AM samples showing the least well defined and least regular AM peak to trough variation. Methods 2 and 3 rely on a clearly and regularly defined periodic variation in the measured data to detect and rate AM. Where this does not arise lower values of AM are attributed. This disagrees with participant's responses. For example, Sample A was given an annoyance rating of 6.4. It contains AM that is intermittent but has AM with peak to trough variation up to 10-12dB.[9] Methods 2 and 3 give this sample an AM value of 3.8 and 4.1 respectively. Thus, AM methods 2 and 3 do not appear to reflect subjective ratings of the samples' annoyance well where the AM is not clearly defined and regular.

4.3 Comments

Much of the value in this study came from the comments that participants provided detailing why some samples of AM were perceived as more or less annoying than other samples and the descriptions of the samples. Table 10 provides a summary of the main descriptors that participants used to describe the samples and an approximate tally of how many times these descriptors were mentioned.

Of interest is the apparent conflict between participants stating that the samples were annoying both due to regularity / constancy and irregularity of the samples. As the comments were optional, clarity as to the meaning of these descriptions could not be sought and lay views may contradict without clear definition. However, it could suggest that both these features cause annoyance equally in participants. It could also be subject to definition, for example whether constant / regular referred to the regularity and constancy of the AM on a second by second basis or that AM occurring in this manner lasting for hours / days at a time would cause annoyance. The rhythmic and repetitive nature of the samples was listed several times as a key annoyance factor and lower frequency sounding AM was generally considered to be more annoying than the higher frequency sounding AM.

A noticeable finding from the comments was that several participants judged some samples to be louder than others. The samples were chosen because they all had the same L_{A90} value and as such should in general terms be judged to be of similar 'loudness' by UK wind farm noise guidance.[10] In reality this may not be the case. Whilst the samples did have slightly different L_{Aeq}'s (see table 1), the levels were not significantly different and varied by a maximum of 3dB, which is generally considered to be a just-perceptible change.[11] This supports an argument that wind farm noise that also contains AM is poorly described using the L_{A90} index and suggests the need for an assessment method for AM that is detached from ETSU-R-97 and its associated L_{A90} values. It is unclear from the results whether the small changes in L_{Aeq} or variation in AM character is driving the perception of different loudness between the samples.

AM was likened to aircraft noise and tumble dryer noise, which is a common likeness identified in earlier research (3). The description of the AM samples having a 'drone' and 'hum' quality was also noted by several participants. This is interesting as such terms are usually associated in acoustics with steady noise sources and those with constant tonal content as sometimes caused by motors. Such

---

[9] This type of AM has been found to be a common feature at most sites where AM has been observed.

[10] Loudness here refers to a lay term and not a formal acoustic definition of loudness, which has not been calculated for these samples.

[11] When in the field rather than in a laboratory environment.

descriptions were responsible for early complaints from wind farms being attributed to low frequency noise rather than AM (17). This description perhaps warrants further investigation alone as to whether participants hear both AM and low frequency noise, whether the combination / interaction of these factors causes complaints or whether it is a miscommunication between those listening to and those assessing wind farm noise.

5. CONCLUSIONS

Numerous methods have been proposed to quantify AM but there are few methods or planning conditions that set a level of acceptability. Limited work has been done investigating whether methods for quantifying and controlling AM reflect subjective judgements of AM. At the outset this paper asked three questions:

- → Does all AM sound the same?
- → Is there a consensus amongst the population as to what is "good" and "bad" sounding AM?
- → Do AM metrics reflect different characteristics of AM?

The MAS online study found that participants do clearly hear differences between samples of AM and that some samples were considered as more or less annoying than other samples. The difference between samples strongly appears a result of character difference rather than differences in energy levels (L_{Aeq} or L_{A90}) as differences in energy levels between samples were small. This has clear implications for methods of rating AM and for those responsible for determining what is or is not acceptable AM.

There was a general consensus that one sample in the study was the most annoying and one sample was the least annoying. Other samples were rated similarly to each other. Whilst there was a consensus over the most / least annoying sample the difference between the annoyance rating of the samples was small. This could suggest either that as soon as AM is present it is generally considered to be annoying and / or that there are multiple characteristics of AM rendering it annoying for different reasons. With regard to comments provided by participants there was some disagreement between features of “good” and “bad” AM. This disagreement mainly focused around features of regularity, irregularity, constancy and repetitiveness. Rhythmic features, low frequency sounds and drones / hums were all listed as contributing to annoyance in accordance with the findings of other research.

Whilst in general the AM rating methods investigated tended to reflect the general trend in annoyance reflected by participants, identifying which samples were more or less annoying, there were limitations to the methods. In particular methods for rating AM significantly underestimated the annoyance response of samples that did not have a regular, clearly defined periodic variation in the measured data. This indicates that caution must be exercised if such an approach is to be taken forward for rating AM as whilst the methods may work well in some circumstances and for some periods of AM they may dismiss many other periods or cases where there is a real AM problem that manifests differently to the ‘ideal’ and regular peak to trough form tested by others.

This paper indicates that all types of AM cannot be treated equally and shows that participants can and do identify the different characteristics in samples of AM. It shows that changes in character other than peak to trough depth affect perception of AM. Although samples of AM were rated differently, the range between ratings was not great suggesting that the mere presence of AM triggers a significant annoyance reaction. Whilst no non modulating wind farm noise was provided against which samples could be compared, this finding was supported by comments provided by participants.

Research has conclusively identified that wind farm noise is more annoying at lower levels than other sources of environmental noise and that the presence of AM is a significant contributor to this annoyance. Further work is needed to identify why AM is annoying, what features of AM contribute to annoyance and whether these factors all contribute to annoyance equally. This work should be used to inform controls for AM rather than having AM controls imposed dictating what listeners should or should not find annoying.

6. FURTHER WORK

This paper presents only the initial findings of the study and further analysis of the results will be completed and reported on at a later date. At this initial stage it is evident that differences between AM samples can be heard and specific and different characteristics are identified between samples. This warrants further investigation in more controlled conditions. Further work in this area should be undertaken and is particularly important given the imminent imposition of AM controls in the UK. Until there is further understanding of the variations in commonly arising AM, the value of any metric addressing one feature of AM cannot be known. Without a better understanding of how those exposed to wind farm noise perceive AM there can be little confidence that controls will adequately resolve justified complaints of adverse impact.

ACKNOWLEDGEMENTS

The author would like to thank all those who took the time to respond to the MAS Wind Farm Annoyance Study. Particular thanks go to the Cotton Farm Residents Association for their continuing support and cooperation in allowing the use and investigation of wind farm noise impact using the data measured with the community permanent monitoring station.

REFERENCES

1. S. Yokoyama, S. Sakamoto and H. Tachibana. Study on the amplitude modulation of wind turbine noise: part 2- Auditory experiments. Proc. Inter-noise 2013, (2013).
2. Renewable UK. Wind Turbine Amplitude Modulation: Research to Improve Understanding as to its Cause and Effect. Available from: http://www.renewableuk.com/download.cfm/docid/528AF1A8-9F36-41E4-8F39042D4BEB795D [Accessed 15 May 2016].
3. Stigwood, M., Large, S., & Stigwood, D. Audible amplitude modulation - results of field measurements and investigations compared to psychoacoustical assessment and theoretical research. Proc 5th International Conference on Wind Turbine Noise; 28-30 August 2013; Denver, Colorado 2014.
4. The Independent Noise Working Group. Wind Turbine Amplitude Modulation and Planning Control Study. Work Package 3.1 - Study of Noise and Amplitude Modulation Complaints Received by Local Planning Authorities in England. Available from: http://www.heatonharris.com/reports-publications [Accessed 15 May 2016].
5. World Health Organisation. Noise and Health. Copenhagen; WHO; 2000.
6. Lee S., Kyutae, K., Choi, W., & Lee, S. Annoyance caused by amplitude modulation of wind turbine noise. Noise Control Eng J. 2011; 59:38-46.
7. The Independent Noise Working Group. Wind Turbine Amplitude Modulation and Planning Control Study. Work Package 5A and 5B - Towards a draft AM condition. Available from: http://www.heatonharris.com/reports-publications [Accessed 15 May 2016].
8. Large, S., & Stigwood, M. The noise characteristics of 'compliant' wind farms that adversely affect its neighbours. Proc Inter-noise 43; 16-19 November 2014, Melbourne, Australia 2014.
9. Institute of Acoustics Amplitude Modulation Working Group. Methods for Rating Amplitude Modulation in Wind Turbine Noise. Available from: http://www.ioa.org.uk/sites/default/files/AM WG%20Discussion%20Document.pdf [Accessed 16 May 2016].
10. A. Fukushima, K. Yamamoto, H. Uchida, S. Sueoka, T. Kobayashi & H. Tachibana. Study on the amplitude modulation of wind turbine noise: part 1-Physical investigation. Proc. Inter-noise 2013, (2013)
11. Cooper, J. & T. Evans. Automated detection and analysis of amplitude modulation at a residence and wind turbine. Proc of Acoustics 2013; 17-20 November 2013; Victor Harbor, Australia 2013.
12. Di Napoli, C. Long Distance Amplitude Modulation of Wind Turbine Noise. Proc Fourth International Meeting on Wind Turbine Noise; 12-14 April 2011; Rome, Italy 2011.
13. Pedersen E, & Waye KP. Perception and annoyance due to wind turbine noise - a dose response relationship. J Acoust Soc Am. 2004;116(6):3460–3470.
14. Van den Berg, F. Wind turbine noise: an overview of acoustical performance and effects on residents. Proc of Acoustics 2013; 17-20 November 2013; Victor Harbor, Australia 2013.
15. Michaud, DS., Keith, SE., Feder, K., Voicescu, SA., Marro, L., Than, J., Guay, M., Bower, T., Denning, A., Lavigne, E., Whelan, C., Janssen, SA., Leroux, T., & van den Berg, F. Personal and situational variables associated with wind turbine noise annoyance. J Acoust Soc Am. 2016;139(3):1455–1466.
16. Berglund, B. & Lindvall, T. (ed.) World Health Organisation, Karolinska Institute (Sweden). Community Noise. Archives of the Center for Sensory Research; 1995; 2(1); 1-195.
17. Hayes McKenzie Partnership Ltd. The Measurement of Low Frequency Noise at Three UK Wind Farms. UK Department of Trade and Industry (DTI). Contract number: W/45/00656/00/00. 2006.Available from: http://webarchive.nationalarchives.gov.uk/20090609003228/http://www.berr.gov.uk/files/file31270.pdf [Accessed 16th May 2016]