A neglected source of uncertainty in potential wind farm noise assessment using the ETSU-R-97 process

By Rod Greenough, Emeritus Reader in Physics, University of Hull, and David Unwin, Emeritus Professor in Geography, Birkbeck University of London

Introduction

The process outlined by the (UK) panel on wind turbine noise in ETSU-R-97 (ETSU-R-97, 1996) has two key inputs, a prediction of the turbine generated noise at selected receptors and survey data on the background noise using the $L_{A90\ 10min}$ weighted measure established over a range of wind speeds referenced to 10m above ground level (AGL). Since its formulation, this assessment process has been criticised and, for better or worse, a suggested improvement, the so-called 'article' method has been widely adopted (see Bowdler, 2006, 2009; Bowdler *et al.*, 2009; Stigwood, 2011; REF, 2012). To date the debate has been on the need to assess the impact of high wind shear on both the extrapolated wind speeds at hub height and the 10m AGL reference height in the height range now spanned by turbines that are significantly larger (now typically 80m AGL) than they were when ETSU-R-97 was defined.

Uncertainties relating to the turbine noise output and manufacturing tolerances (Broneske, 2009), the assumed ground absorption, atmospheric attenuation, the accuracy and resolution of the sound recording instruments, and their ability to filter true background from noise induced by the wind itself have also been considered and the related uncertainties in noise margins (the difference between the predicted noise at a receptor site and the allowed noise level according to ETSU) will be analysed in a future publication.

Our principal concern here is prompted by a comparison of several wind farm applications in which the applicants claim, correctly, that the ETSU-R-97 regulations have been adhered to. The problem arises when the recommended procedures for the analysis of measured data reach the stage when the onus is on the applicants to adopt reasonable and meaningful analytical methods. Without employing models based on well-established data analysis and statistical techniques, each applicant performs regression analysis as a basis to determine the allowed ETSU noise levels. After surveying many applications, it is evident that there is a marked lack of consistency in these analyses. It is this source of uncertainty, which arises from the models used by applicants in the establishment of an average background noise curve as a function of the 10m AGL wind (V_{10}), which is addressed in this note.

Background: ETSU-R-97and the background polynomials

ETSU-R-97 (page 101) outlines how the panel expected background curves for noise to be obtained as follows:

"For each sub-set, a "best fit" curve should be fitted to the data using a least squares approach, usually a polynomial model (of no more than 4th order). Where there is considerable scatter in the data, it may be more appropriate to bin the acoustic data into 1m/s bins before identifying a best fit model. These two curves, referred to as the 'day-time curve' and the 'night-time curve', provide a characterisation of the prevailing background noise level for day-and-night respectively, as function of wind speed from zero to 12m/s at 10m height. Note that whatever model is used to describe the measured data, this should not be extrapolated outside the range of the measured wind speed data."

Further we are also told that:

"The variation in background noise level with wind speed will be determined by correlating $L_{a90,10min}$ noise measurements taken over a period of time with the average wind speeds measured over the same 10-minute periods and then fitting a curve to these data."

The ETSU-R-97 advice most frequently followed is to fit a best fit polynomial curve to the background noise data using the standard 'ordinary least squares' (OLS) criterion of fit, which under some well-understood assumptions provides the best linear unbiased estimates for the coefficients that define this curve. These fits have the general form $Y = F(x)$ in which Y is the background sound level in dBA (ten-minute average) at a neighbouring dwelling's amenity area, x the measured or inferred wind speed at 10m AGL (V_{10}) at the wind turbine site, and F denotes 'some function'. It would seem that the ETSU-R-97 panel were of the opinion that specification of a polynomial of up to the 4th degree for F(x), coupled with the use of the phrase 'best fit' were sufficient to ensure a reasonably objective and robust result on which the planning process could rely. Fits to the observed data are usually reported using the *coefficient of determination*, or R^2, a statistic that is probably better thought of as the *percentage of the variance explained* by the fitted curve. These curves are what here we call *models* of the underlying data, but the guidance says very little about why these quite complex polynomials have been used, or any caveats that should perhaps be attached to them, yet the establishment of a reliable curve for the background noise is critical for determining noise impact on neighbouring dwellings and setting fair noise conditions to protect amenity.

At the outset, it is worth commenting on several statistical issues that arise from this approach:

- The coefficients arrived at by so-called ordinary least squares (OLS) multiple regression are themselves *estimates* of some *unknown parameters* in the full population from which the sample background ($L_{A90\ 10min}$) and wind (x, V_{10}) were sampled and as such are themselves subject to an uncertainty that should be expressed as a confidence interval around the plotted line;
- ETSU-R-97 assumes that the main driver for the observed variation in background is wind speed and it is utterly reliant on these plots and fitted functions. We have yet to read a justification for the implied correlation either in theory or by means of careful measurement at proper free field locations using correctly shielded ground level microphones. At some sites the major cause of variation in background might well be some other process of which the regular hum of traffic close to a main road is probably the most important example. Background in such cases would correlate more closely with time of day and wind direction than with wind speed;
- The 'explained variance' given by the R^2 value refers to a statistical notion of an 'explanation' that should not necessarily be equated with scientific causation;
- Although we are advised that polynomials of degree higher than four should not be used, this is without any additional comment or justification and fifth order polynomial fits are not unknown;
- A major failing of ETSU-R-97 lies in the way that the measured data are assumed to be unproblematic. They are not. Typically, an Environmental Impact Statement (EIS) required by the Local Planning Authority (LPA) will have an assessment of the likely noise nuisance at selected receptors for both 'quiet daytime' and 'night-time' conditions based on the established curve of background noise plotted against V_{10} winds using observations collected over at most a few weeks simultaneous recording of both $L_{A90\ 10min}$ (dBA) at the receptors and V_{10} (m/s) at the

wind farm site and either inferred from the wind profile at a high mast or measured using a meteorologically standard 10m mast. The uncertainties related to how the V10 derived from a mast are 'standardised' have been well documented (Bowdler, 2009), but what is often forgotten is that these data are a usually a very poor sample in both time and space. In time they are a snapshot of background noise for a very limited portion of the year, an issue that REF (2012) demonstrate could introduce +/- 5dBA difference, and hence uncertainty, in the fitted curves;

- In space, reliance on V_{10} measure at a single point in what typically will be a moderately large area of possibly highly spatially varying wind regime introduces even more uncertainty that has yet to be quantified. Moreover, contamination of the data by transients will frequently occur and the possible influence of wind induced noise at inadequately shielded microphones has yet to be resolved, giving yet more uncertainty;
- What is almost always forgotten is that this curve fitting procedure, using classical regression, that has been known and used since the mid-nineteenth century, assumes that the data are an independent random sample from a defined population of possible values. The method evolved when, rather than being a very large data file downloaded from an automatic recording device, each and every data point was likely to be hard won by careful hand measurement;
- Both numbers, the background and the reference wind speed, come from a time series sampled over ten-minute intervals. It is inevitable that such data will to a greater or lesser extent exhibit *auto-* or *self- correlation.* Autocorrelation can be understood by a simple thought experiment. Suppose that at some time the anemometer records a V_{10} of 10m/s, what is the value likely to be in ten minutes time? Given that meteorological elements show persistence in time it is highly unlikely to be either 0 m/s or, say, 25m/s. Chances are that it will be fairly close to 10 m/s. In other words successive data are correlated with themselves. Yet statistical inference assumes that each case is independent or uncorrelated with the others. The effect on the result is to bias the standard error because the standard goodness of fit measures are tricked into believing that there is a larger sample than actually exists. Larger samples give smaller standard errors and better statistical significance;
- Finally, the number of sample points (n) is not only large but is to a very large extent *arbitrary*; it can be almost as large as the analyst likes (for example by using more weeks data, or decreasing the sampling time interval), but the impact on the statistical significance of any results is to make any change, not matter how small, almost certain to pass the standard tests. There is a real risk here of conflating the statistical notion of 'significance' with the scientific one and it cannot be stressed too highly that they are not the same thing.

This is not the place to enter into a long exegesis of the assumptions of linear regression and their impacts on the fitted curves, nor do we argue for complete statistical purism: there are literally millions of successful scientific studies that at some point break one or more of these assumptions.

What we should point out is that regression was introduced as a means by which specific scientific hypotheses, for example those generated from physical reasoning, could be tested and/or calibrated against observation of the real world. The ESU-R-97 document and hence the process it mandates says absolutely nothing about the underlying physics of wind generated noise. At no point in the ETSU document or anywhere else in the literature, can we find any physical justification in physics, acoustics or meteorology for the choice of model to be fitted. This has some serious consequences for the reliability of the entire process.

Any background line will do?

In all the environmental impact statements (EIS) associated with wind farm noise assessments we have examined what we find are polynomial curves of degree $p = 2$ (quadratic), $p = 3$ (cubic), sometimes $p= 4$ (quartic), and in one case even a degree $p = 5$ (quintic) fitted to the background and wind data. The occasional commentary in the text shows that the fitting process seems almost always to be driven by an obsession with the idea *best fit* being equivalent to 'highest coefficient of determination, R^2, I can get'. Table 1 illustrates the uncertainties this model choice introduces into an assessment with results from various equations used in the analysis of data (825 data points) from a recent wind farm case.

Type of fit	Equation	R^2
Linear p=1	$y = 1.7655x + 21.011$	0.59
Quadratic p=2	$y = 0.0312x^2 +1.3689x + 22.081$	0.59
Cubic p=3	$y = 0.0289 x^3 - 0.5175 x^2 + 4.3399x + 18.243$	0.60
Quartic p=4	$y = -0.0049x^4 + 0.1585x^3-1.6772x^2 + 8.2981x +14.558$	0.61
Quintic p=5	$y = -0.0028 x^5 + 0.0853x^4 - 0.9268x^3 + 4.1847x^2 - 5.2866 x + 24.408$	0.62
Exponential	$y = 22.329 e^{0.0555x}$	0.61

Table 1: Results from various model fits to background noise data and the corresponding regression coefficients In these equation Y is the dependent variable $L_{A90\ 10min}$ (dBA) and the independent variable x is the inferred wind speed at 10m AGL (m/s)

The polynomials of degree $p = 2$ or $p = 3$ are those that almost certainly would have been accepted as appropriate models on which to base the ETSU assessment, but we cannot resist pointing out that an alternative, equally plausible, model that actually fits the data better than all but the degree $p = 5$ polynomial is the rather elegant exponential.

Unless this is to be a scientific hall of mirrors, which of these models should be used in the assessment or will any curve do the job just as well? All suggest that with no wind the background is somewhere between $L_{A90\ 10\ min} = 14.558$ and 24.408 dBA, which seems reasonable for a quiet rural location, and all describe the data reasonably well, giving coefficients of determination in the range $R^2 = 0.59 – 0.62$. We suspect that, faced with this choice and secure in the knowledge that almost every planning decision maker would accept their 'professional judgment', it would be a brave acoustics consultant who did not chose the model that best suited their employer's objectives but statistical analysis and physical logic can help a little in this choice.

One formal statistical option can be understood by the observation that straight line, degree $p = 1$, polynomial requires estimation of two coefficients whereas the degree $p = 5$ quintic one requires estimation of 6 for a gain in 'explanation' in the above example of just 3% (=100 x (0.62-0.59)). Statistically speaking, there is a clear case here for an appeal to Occam's Razor P38▶

◀P37 suggesting that the simplest model that is consistent with the data is the one that should be fitted. As polynomials of progressively higher degree are used they allow the curve to add points of inflexion around which it can twist to accommodate the observed data. It is inevitable that this added flexibility will increase the R^2 and so in some sense be a 'better fit', but the danger is that of *over-fitting*, introducing features into the curve that are artefacts solely of the degree of function chosen (p) and have nothing to do with nature itself. It follows that the statistical question that should be asked is NOT '*is this new model of degree p+1 a better fit to the data than the model of degree p*', but '*given that we have to estimate another coefficient, does this new model of degree p+1 significantly improve on the fit given by the model of degree p?*' This is a question well known in data analysis in general and specifically to geostatisticians in the context of fitting polynomial regressions, called trend surfaces, to the locational coordinates of mapped information (see for example O'Sullivan and Unwin, 2010, pages 279-287) and a simple analysis of variance approach has been adopted to handle it. Applying this approach to this case, what we find is that, even with such a large number of data points, the addition of the quadratic is only just significant at the 95% level (i.e. one chance in twenty of being wrong) , but not at 99%. Similarly the very large, n, of strongly autocorrelated data points made available by courtesy of the recording devices, ensures that the cubic and higher order terms are also just statistically significant, but almost any statistician confronted with these results would counsel caution and warn against over-fitting.

It should be stressed that in standard noise assessments any of these models could have been presented, accepted as definitive, and used to set what would have been asserted to be ETSU-R-97 compliant limits. Much of the difficulty that the approach defined in ETSU-R-97 generates could be avoided by making it clear that this step is one of model *selection* in which the objective is to choose the model that gives the best predictions from a range of possibilities. As computing power has increased, modern statisticians have developed a number of strategies and measures for precisely this purpose. Of these the Akaike Information Criterion (see Akaike, 1974), which combines a measure of the model fit with a penalty related to the number of parameters that have to be estimated, is the best known and most widely used.

Using other regression diagnostics?

There are alternative ways of fitting curves to plots and there are alternative regression diagnostics to the crude R^2 coefficient of determination. Using a simple statistics package there is often the facility to identify *unusual observations* that are either badly fitted or that exercise undue influence (called their *leverage*, see Unwin & Wrigley, 1987). Of interest in the context of model selection is the distribution of *unusual observations*, something that is not necessarily apparent from a visual examination of the plotted line and the scatter of data points

For the linear fit, degree $p = 1$ polynomial in the example from Table 1, the software we have used (MINITAB) identifies 86 unusual observations of which 32 are badly fitted having a high *standardised residual* (the value divided by its standard deviation) and 54 have undue *influence* on the fitted line indicated by a high *leverage*. Of the badly fitted points the majority (24 from 32) have negative residuals. Of rather more significance to our argument are the 54 observations that exert undue leverage on the solution. Leverage is also known by the phrase 'distance to the centre of the data' and in the example this is very evident, but with a particular bias towards observations at low winds. In fact 52 of these points are at V_{10} winds less than 2.0m/s which leads directly to a very important point of principle: *although most assessments might choose to ignore the data at low winds less than 'cut in' of the turbine, these data have disproportionate importance in 'fixing' the shape of the model fitted to the entire data set.* In fact, the behaviour of the model close to the V_{10}=0, no wind, axis is critical. This is unfortunate, not least because in such very light air cup-based anemometry is not very reliable and there may be issues relating to the calibration, zeroing, and possible drift of the instruments used.

As can be seen from an examination of the estimated coefficients, and the similarity in R^2, in Table 1 fitting the quadratic makes very little difference and the same issues emerge. In this case 89 observations are identified of which 30 are badly fitted and 59 now have undue influence on the fitted line indicated by a high *leverage*. In passing, note that reliance on the linear curve gives the possibility of departures at some time or other of up to +/- 10dBA which is a doubling or halving of the predicted sound level from the curve

Appeals to logic?

We have already noted that, in the seeming absence of any theoretical expected forms for these curves there is clearly a blind reliance on getting a good fit as measured by the *coefficient of determination, R^2*. However, even without the benefit of acoustic theory, we can make some progress by appeals to simple logic and can illustrate this by a sequence of no less than three models offered in response to various objections at another recent public inquiry, again for the quiet daytime at an obviously at-risk receptor.

The initial attempt, shown here as Figure 1, used a simple degree p= 2 quadratic with a plot showing all of the data down to close to V_{10}=0 m/s, but did not report the R^2:

Degree p= 2: $L_{A90\ 10min} = Y = 0.0587x^2 + 0.5167x + 30.548$ dBA

Note that this suggests an arguably high background in a very quiet rural area at V_{10}=0 m/s of $L_{A90\ 10min}$ = 30.548dBA. Responding to a query from the local environmental health officer, the next attempt used a different method of referencing the winds to 10m AGL and some additional survey data to produce the degree p=3 cubic model shown as Figure 2:

Degree p= 3: $L_{a90\ 10min} = Y = -0.0513x^3 + 1.1815x^2 - 6.1697x + 39.99$ dBA (R^2= 0.5551)

P40▶

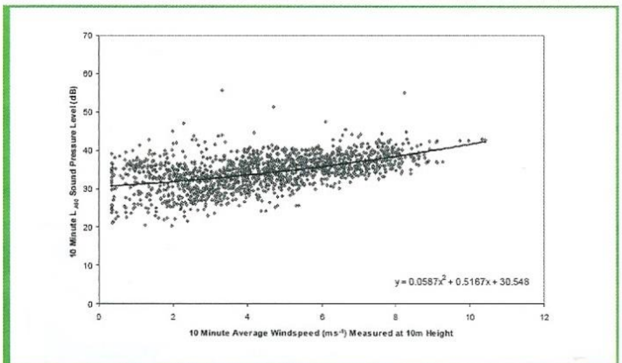

Figure 1: The quadratic model

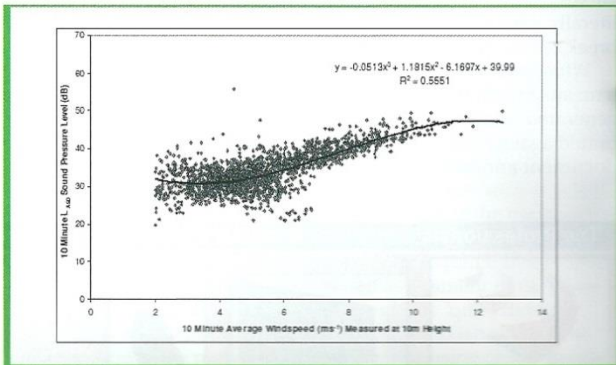

Figure 2: The cubic model

◄P38 At least this reports the rather modest fit that is inevitably higher than that for the quadratic, but at what cost in logic do we get this improvement? Notice that the introduction of a cubic term (x^3) into the equation means that we now allow the function to have two points of inflection at which its curvature changes from being concave upward (positive curvature) at low wind speeds to concave downwards (negative curvature) at higher speeds. What matters here isn't whether or not the additional term significantly improves the fit but whether or not it makes sense in simple logic. It does not.

Many wind farm noise assessments argue that below the cut in speed of the turbines (say 4m/s at hub height) the shape of these curves is not important; in fact the behaviour of the function used as it approaches and meets the background noise vertical Y-axis is critical. There are two important physical considerations. First, the intercept at the Y-axis represents the background noise at any chosen site in the absence of wind. Logically, and from simple physical considerations, when there is no wind *we would expect similar geographical locations scattered around the wind farm turbines in the same area to have consistently similar values for background noise.* Second, *we would expect the curve to flatten steadily towards the same axis and, to have a zero gradient where it meets the axis.*

Neither of these conditions is met in the example shown in Figure 2. First, at V_{10}= 0 m/s it predicts a background in the quiet daytime hours at a site in a very quiet rural area of an extremely unlikely $L_{A90\ 10min}$ = 39.99dBA. Second, although the full extent of this feature is hidden by the 'blanking out' on the plot of many of these data from V_{10} = 0 to around V_{10} = 3 m/s, it suggests that as the wind increases so the background noise gets less, which is equally unlikely. In our opinion both features, the high intercept and the negative gradient, have nothing to do with nature and everything to do with over-fitting a cubic model to data that do not warrant it. Any cubic function will inevitably bend through two points of inflection and that it is inevitable that this extra freedom for bend will increase the goodness of fit as measured by the R^2. If a cubic function fitted by least squares doesn't show two points of inflection in the range of the data, logically it must be the wrong function: a quadratic would have done the job just as well. Finally, at a late stage in the planning process a third model that attempted to correct some of these problems was offered and is shown in Figure 3.

Degree p= 3: $L_{A90\ 10min} = Y = -0.021x^3 + 0.4936x^2 - 1.7502x + 31.703$ dBA (R^2= 0.6766)

This has the same cubic shape as before and a better fit. Other than the use of a properly estimated V_{10} wind and the fact that the correlation seems to be improved we are not told anything more about how it was derived. It can be seen that it removes all the data for V_{10} speed below 2m/s so concealing the fact that once again we have a negative gradient in this range. At $L_{A90\ 10min}$ = 31.703 dBA the background at V_{10}=0m/s once again appears on the high side.

Does it matter?

Does it matter that in the range of wind speeds that are of concern that we have different versions of the background curve that the ETSU-R-97 process requires? For the various models listed in Table 1, at V_{10} = 5m/s the background curve value to be used in the assessment is as is given as in Table 2

Model fitted	Background at V_{10}=5m/s $L_{A90\ 10min}$ (dB)
Polynomial, degree 1	25.35
Polynomial, degree 2	24.48
Polynomial, degree 3	24.10
Polynomial, degree 4	23.99

Table 2: Background noise $L_{A90\ 10min}$ (dBA), at V10=5m/s Case 1

For the models presented in our second example in Section (5) the equivalent background values are as in Table 3.

Model fitted	Background at V_{10}=5m/s $L_{A90\ 10min}$ (dB)
Initial Quadratic	34.60
Polynomial Degree 3, Model (2)	32.27
Polynomial Degree 3, Model (3)	32.67

Table 3: Background noise $L_{A90\ 10min}$ (dB), at V10=5m/s Case 2

In both cases even at V_{10}=5.0m/s there is a range of background values of around 1.4 - 2.0dBA in the $L_{A90\ 10min}$, which increases at V_{10} lower than this and decreases as V_{10} increases above it. This range has very little to do with nature and everything to do with the choice of model fitted to the data. The uncertainty is less than that reported as arising with different corrections for wind shear (Stigwood, 2011) and, although modest, it could well be important in any decision made with receptor sites that are marginal in the ETSU-R-97guidance.

It should be stressed that *any* of these curves could well have been used in determination of an application to build a wind farm and/or in the determination of critical limits for related conditions. That any one or other of them increases or decreases the reference values at the receptor sites, and so does or does not favour a developer, is in our opinion irrelevant. Just as by manipulation a developer might be able to raise the background by choice of data and function, so could any competent data analyst find a function that would lower it by the same, or even greater, amount. The difference is that an honest data analyst would be well aware of this fact, report the uncertainty, and suggest allowing for it in any decisions based on it.

Is there an alternative?

Given these scatter plots, there are at least three alternative ways of reaching the representative values on which ETSU-R-97 relies.

(i) *Locally Weighted Scatterplot Smoothing*

First, professional statisticians would undoubtedly suggest alternative ways of fitting and assessing the fit that would address issues of model choice and the 'messy' character of the data. Of these the most obvious is *locally weighted scatterplot smoothing or local regression* (LOESS), that fits local models that derive their form from the data themselves rather than having to be specified *a priori* by the analyst is an approach that is widely used to isolate the 'signal' from the 'noise' in this type of plot and (see for example Cleveland, Grosse and Shyu, 1992). This type of smoothing is available in several software packages, but it relies on the user supplying a parameter that controls the degree to which the data are smoothed and so is open to possible manipulation by the analyst.

(ii) *Direct use of mean values with 'binned' data* ►

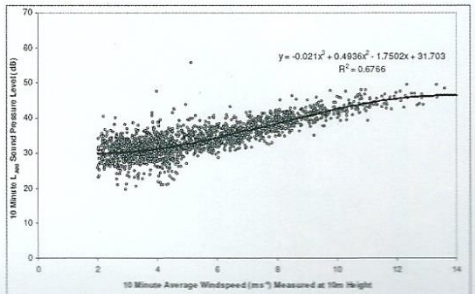

Figure 3: A second cubic model

Case Number (Day/Night)	Background at Zero wind constrained Method $L_{a90,\ 10min}$ (dB)	R^2 constrained Method (%)
Day 1	23.5	72
Day 2	20.9	58
Day 3	22.0	62
Day 4	23.4	73
Day 5	21.6	64
Day 6	23.4	73
Night 1	17.4	82
Night 2	16.5	69
Night 3	15.8	74
Night 4	17.7	80
Night 5	19.4	64
Night 6	19.8	70

Table 4: Summary values for polynomial of degree p=4 fitted to six different receptors (1-6) for both day and night time conditions in a wind farm noise assessment from eastern England.

Second, and much more transparently, referring back to the original ETSU-R-97 recommendations we find a sentence (page 101) that indicates that the panel were aware of a simpler alternative, which is to smooth the data before undertaking the regression analysis:

Where there is considerable scatter in the data, it may be more appropriate to bin the acoustic data into 1m/s bins before identifying a best fit model.

For reasons that we do not understand, this simple option seems subsequently to have been totally ignored and in fact there is no need whatsoever to undertake any regression analysis. Figure 4 shows a summary of the data used to prepare Table 1 ' binned' at the nearest whole number wind speeds and presented as a sequence of stacked box plots.

In each graphic the vertical line shows the total range of the data in each bin whilst the rectangle shows the inter-quartile range and the horizontal line is at the median value for that bin. This display has the merit of showing the very considerable scatter that exists around any measure of the central tendency in each wind speed bin. In every example we have examined such a display would have been sufficient, and there is no need to go further and use regression analysis on the binned means or medians, but if there is an insistence on finding a 'best fit' function the most appropriate shape seems obvious. Binning data in this way has two disadvantages. First, it literally 'throws away' information that could be of use and, secondly, it introduces a dependence on the arbitrary boundaries of the bins. A clear advantage is that, although it might make the choice of function and variation around that function easier, each of the bins can be carried forward, complete with their individual gauges of uncertainty, for incorporation in the ETSU-R-97 assessment of noise margins without any need to fit a function. Indeed, presentation of boxplots for each integer wind speed together with the predicted wind turbine noise on the same graph would have the great merit of showing how safe the allowed headroom in ETSU-R-97 would be for each and every receptor and time period.

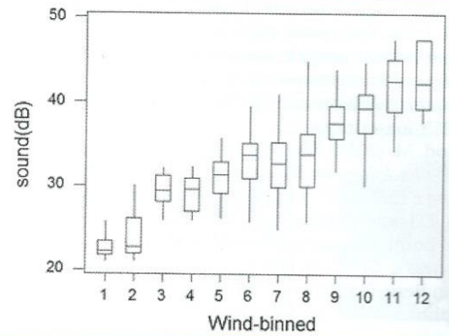

Figure 4: Stacked box plots at each whole number wind speed

(iii) *The zero-gradient at the Y axis approach*

In Section 5 we note that a simple constraint on the fitted curve is provided by the observation that at the point it intersects the background noise (vertical) axis, the rate of change of noise with wind speed must be zero. This constraint is easy to apply if we rely on polynomials of degree that are an even number, in practice either a quadratic (p=2) or quartic (p=4).

Figure 5 shows results from data typical of background noise as a function of wind speed. In the first plot is a conventional quartic (p=4) polynomials fitted to these data. This is followed by a second plot using a quartic function constrained to cross the Y-axis with zero gradient. We would argue that this is a much more plausible curve for these data which also gives a zero wind background of just 17.5dBA. The loss of fit, as measured by the R^2, is negligible.

The advantages of this approach can be illustrated by comparing it to the conventional approach for six 'at risk' receptors at a proposed site in eastern England with the results as shown Table 4 for the constrained fits using exactly the same data as measured and used by the applicant for each receptor site. In every case the function fitted was a quartic, p=4, polynomial.

There are a number of features of note in the comparison of the conventional and constrained results. First, according to the applicants ES, using the standard method without any P42

◄P41 gradient constraint, the applicant's curves all clearly overlay the daytime data points just as the standard method illustrates in Figure 5. However they generate wildly varying values for the background noise at zero wind that span a range of 50dBA and include physically impossible negative values. The night time values using the standard approach are much more stable covering a range of 4dBA but seem inappropriately high for what is a very quiet rural location. As shown in Table 4, imposition of a zero gradient at the Y-axis constraint has a marked and welcome effect on both day and night time zero wind speed background noise values at all six sites. There are several consequences. During the day the effect is to stabilise them in the range 20.9 to 23.5dBA which, given the similarity of the locations, is much more plausible. At night the same effect is seen, but now there is a reduction to a barely measureable background of 16.5 to 19.8dBA. Although in every case the constrained curve data R2 shown in the final column will be necessarily less than that obtained with the unconstrained method, this reduction in value is at most only 1 to 2%. Our view is that, by its use of simple physical reasoning, this is the best of the three suggested options in this Section and for all the cases we have examined, this easy approach to the curve fitting process provides much more consistent estimates of the 'zero wind' background noise and a shape of curve that is in accord with common experience.

Conclusion

In conclusion, we note that the variation in the fitted curves and their impact on the values taken as representative of the background noise at each and every at risk receptor generates a neglected, but very real, *uncertainty* in the entire ETSU-R-97 process. We have demonstrated that replacing the blanket recommendation that a 'best fit' polynomial curve should be fitted to summarise these data before comparison with the predicted turbine noise by either a simple locally weighted average smoothing, a smoothing using the already binned data, a simple set of boxplots of these binned data without any accompanying function, or a curved constrained to intersect the Y-axis at zero gradient results in a reduction in this uncertainty.

Our analysis does not of course include other uncertainties related to the time period of the sampling of the sound data, calibration and related instrumental errors in the meters used, the type of turbines to be installed, variation in sound output from nominally the same machinery, the noise prediction methodology adopted (especially the allowance for ground absorption and/or reflection), and the way that both the wind at hub height and 10m AGL are adjusted to allow for the continuously variable wind shear. Even quite modest estimates of all these uncertainties suggests 'worst case' scenarios that could easily double or halve the background at a receptor. Given that wind farm consents are routinely given with headroom values in the operational range of the turbines of a few decibels and do not recognise the uncertainties that surround the estimates used, it seems inevitable that breaches of any imposed planning conditions will occur. Since we know of no case where noise nuisance has resulted in a consent being denied, the reverse, that consents are being denied when the same considerations of uncertainty should suggest the reverse, does not apply.

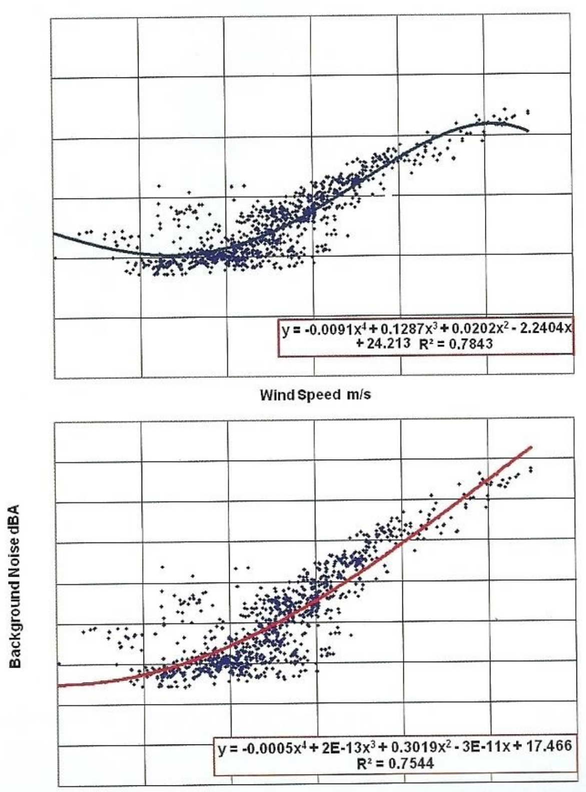

Figure 5: The standard ETSU-R-97 suggested approach compared with application of a zero gradient constraint using typical background noise data

References

1. Akaike, H. (1974) A new look at statistical model identification. IEEE Transactions on Automatic Control 19, 716–723.
2. Bowdler, D. (2006) *ETSU-R-97: Why it is wrong.* Dick Bowdler, Acoustic Consultant, 8 pages
3. Bowdler, D. (2009) Wind shear and its effect on noise assessment of wind turbines. Dick Bowdler, Acoustic Consultant, 14 pages
4. Bowdler, D. , Bullmore, A., David, B., Hayes, M., Jiggins, M., Leventhal, G. and A. McKenzie. (2009) Prediction and assessment of wind turbine noise. *Acoustics Bulletin*, 34(2): 35-37
5. Broneske, S. (2009) *Comparison of wind turbine manufacturer's noise data for use in wind farm assessments.* Paper presented at Third International Meeting on Wind Turbine Noise, Aalborg, Denmark, 17-19 June 2009, 10 pages.
6. Cleveland, W.S., Grosse, E. and M. J. Shyu (1992). Local Regression Models In J. M. Chambers and T. Hastie, (eds) *Statistical Models in S*, (New York, Chapman and Hall), pages 309-376.
7. Cox, R, Unwin D & T. Sherman, (2012) *Wind Farm Noise Assessment: Where ETSU is silent* (80 pages)
8. Davis, J.C. (2002) *Statistics and Data Analysis in Geology*, 3RD Edition pages 407-411 (Wiley: Chichester and New York)
9. ETSU-R-97(1996) *The Assessment and Rating of Noise from Wind Farms*, (The Working Group on Noise from Wind Turbines, Harwell, Oxon), 150 pages
10. Hayes McKenzie Partnership (2011) *Analysis of How Noise Impacts are considered in the Determination of Wind Farm Planning Applications*, HM 2293/R1
11. O'Sullivan, D. and Unwin, D.J. (2010) *Geographic Information Analysis* (New York, John Wiley and Sons, Second Edition)
12. REF Renewable Energy Foundation (2012) *A Critique of the IoA Treatment of Background Noise for Wind Farm Noise Assessments* Available at http://www.ref.org.uk/publications/255-ioa-critique
13. Stigwood, M. (2011) *The effect of a common wind shear adjustment methodology on the assessment of wind farms when applying ETSU-R-97.* MAS Environmental, 56 pages
14. Unwin, D.J. and Wrigley, N. (1987) Towards a general theory of control point distribution effects in trend surface models, *Computers & Geosciences*, 13: 351 355.
15. Van den Berg, G.P. (2006) Wind-induced noise in a screened microphone. *J. Acoust. Soc. Am.*, 119: 824-833.