Estimating mixture models for environmental noise assessment

Gordon Hughes

School of Economics
University of Edinburgh

7th September 2017

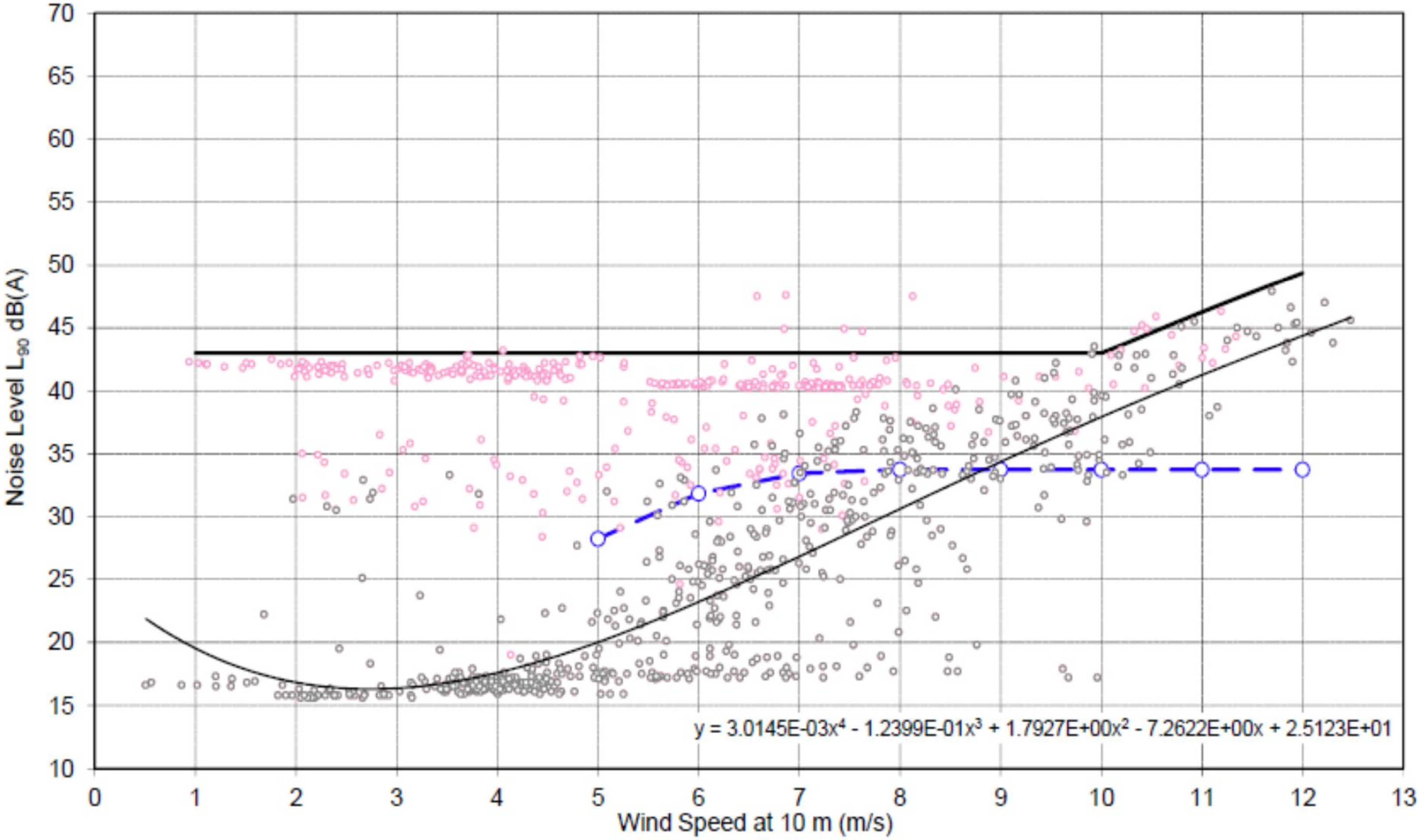
What is environmental noise assessment?

- Assess the impact of a project/development on noise experienced by people/properties nearby
  - Factory, road, airport, power plant, wind farm
- Need to establish a pre-project baseline – background noise level – excluding similar nearby sources
- Estimate expected level of noise from new source at the receptor and potential consequences
  - Loss of amenity, damage to health (esp sleep disturbance), etc
  - Likelihood of complaints about statutory nuisance
- Establish operating noise limits – both output & exposure
  - Where necessary identify measures to mitigate impacts
- Key issue: what is the background noise level?

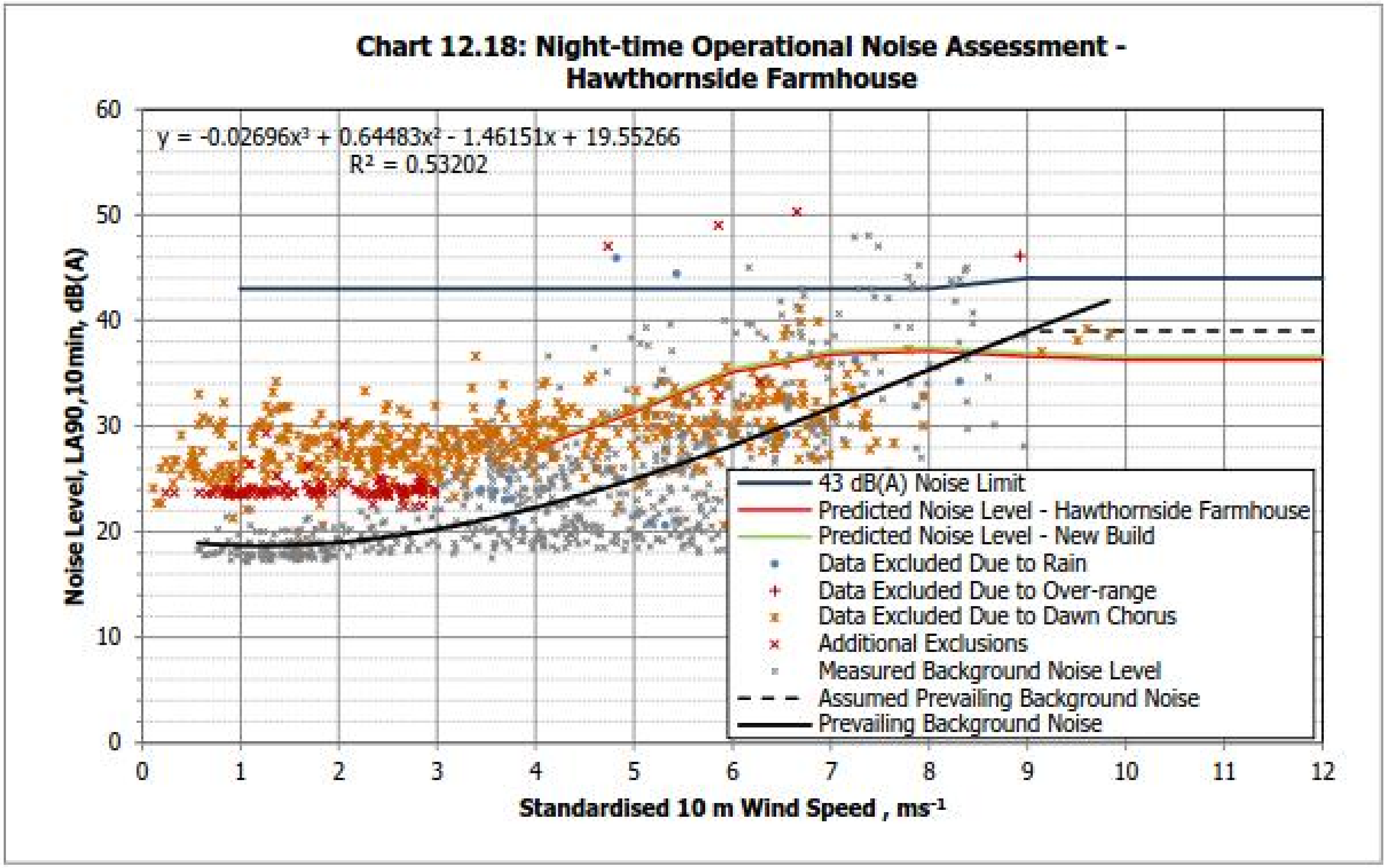
Nature of noise assessment data

- Noise levels measured over 2-4 weeks
  - Covariates – wind, rain, time of day
  - Weighted average of different frequencies
  - LA90 is 10th percentile of continuously measured values for each 10 min reporting period
  - Reported in decibels – log10 scale – but physical relationships relate to sound power – $10^{(dB/10)}$
- Statistical methods used to estimate background noise
  - No clear definition but it is interpreted as the ambient noise excluding (a) the source(s) to be analysed, and (b) intrusive intermittent peaks in the noise level
  - Usual focus is on quiet day-time or night hours

Examples of 'good practice' 1



Examples of 'good practice' 2



What are the problems with standard practice?

- Exclusion of supposed outliers – why & how?
  - What is the data generation process?
- Use of standardised wind speed at 10 metres
  - Measurement error – biased coefficients
  - Not location-specific – better to use wind speed at hub height
- Why fit a random polynomial to decibels?
  - Additive errors implies use of sound power not dB
  - Data suggests some kind of threshold in wind speed (~4-5 m/s)
- Distribution of the LA90 as an extreme order statistic
  - Some assumptions imply use of a beta distribution
  - More generally, use a flexible distribution with positive skewness – consider either the lognormal or the Weibull

Finite mixtures model

Noise N_{jt} measured at location j in period t:

$$N_{jt} = \sum_{m=1}^{M} d_{jt}^{m} \varphi(\mathbf{Z}_t \boldsymbol{\beta}_j^m, \sigma_j^m)$$

where m = 1…M denotes the mixture number, $d_{jt}^{m} = 1$ if the measurement for period t at location j is drawn from component m and is zero otherwise, $\varphi(\mu, \sigma)$ is an appropriate distribution with a location parameter of μ and a scale parameter σ. The location parameter is expressed as a linear function of Z_t, a vector of covariates. The likelihood function is formulated in terms of the probabilities $p_{jt}^{1} \ldots p_{jt}^{M}$ that an observation for location j in period t may be drawn from each of the component distributions, which may be conditional upon the values of a set of covariates.

Implementing the specification

- Beta distribution requires (0,1) variables
  - Transform noise using range (min-5, max+5)
  - AIC/BIC values adjusted to allow for transformation
- Covariates – wind speed at hub height, night/quiet day, cumulative rainfall over 2h & 24h (stream noise)
- Filtering data to remove contribution of other wind farms
  - Difficult and contentious – produces datasets that are prone to non-convergence
- How should background noise be defined in this context?
- Is it worth pooling data using panel methods?

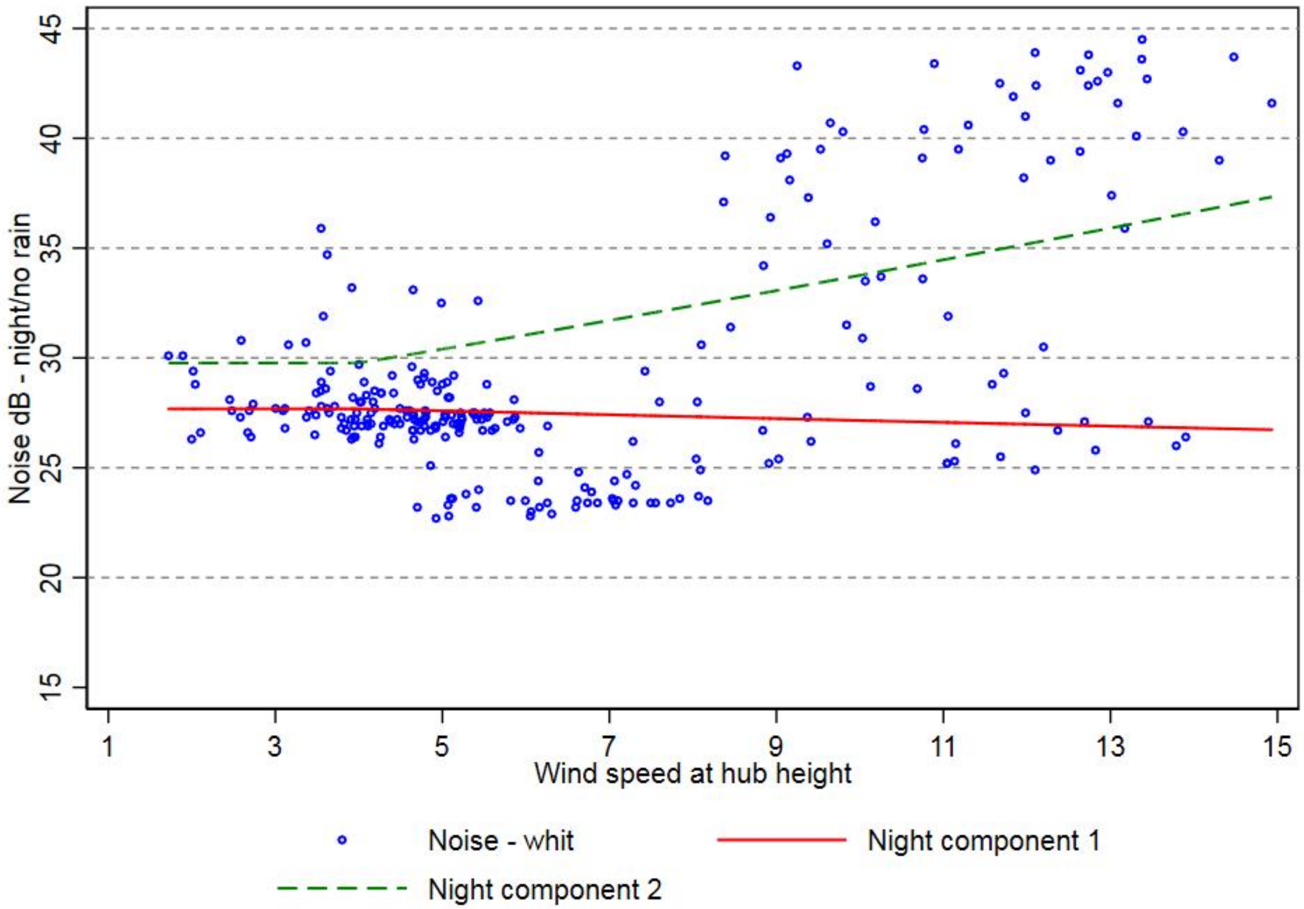
Model comparisons using BIC – location 1

Site	Components	BIC for model: Beta	Normal	Lognormal	Weibull
1	1	11,855	11,961	11,851	12,328
1	2	11,569	11,569	11,610	11,750
1	2 R	11,564	11,561	11,603	11,775
1	3	11,464	11,456	11,521	11,674
1	3 R	11,454	11,451	11,528	11,694
1	4	11,441	11,489	11,499	11,671
2	1	12,609	12,764	12,637	13,029
2	2	12,336	12,327	12,357	12,490
2	2 R	12,347	12,350	12,396	12,543
2	3	12,311	12,304	12,330	12,379
2	3 R	12,307	12,305	12,324	12,476
2	4	12,314	12,323	12,330	12,383
3	1	13,347	13,561	13,739	14,492
3	2	12,890	13,184	13,137	13,767
3	2 R	13,058	13,457	13,374	13,986
3	3	12,805	12,960	12,878	13,435
3	3 R	12,907	13,109	13,094	13,721
3	4	12,808	12,911	12,826	13,268
4	1	12,066	12,129	12,007	12,228
4	2	11,614	11,608	11,654	11,783
4	2 R	11,614	11,606	11,654	11,778
4	3	11,638	11,592	11,658	11,727
4	3 R	11,604	11,581	11,640	11,712
4	4	11,612	11,606	11,636	11,700
5	1	13,585	13,884	13,630	14,051
5	2	13,329	13,413	13,348	13,563
5	2 R	13,348	13,471	13,378	13,644
5	3	13,294	13,299	13,286	13,461
5	3 R	13,317	13,346	13,324	13,544
5	4	13,287	13,318	13,293	13,390

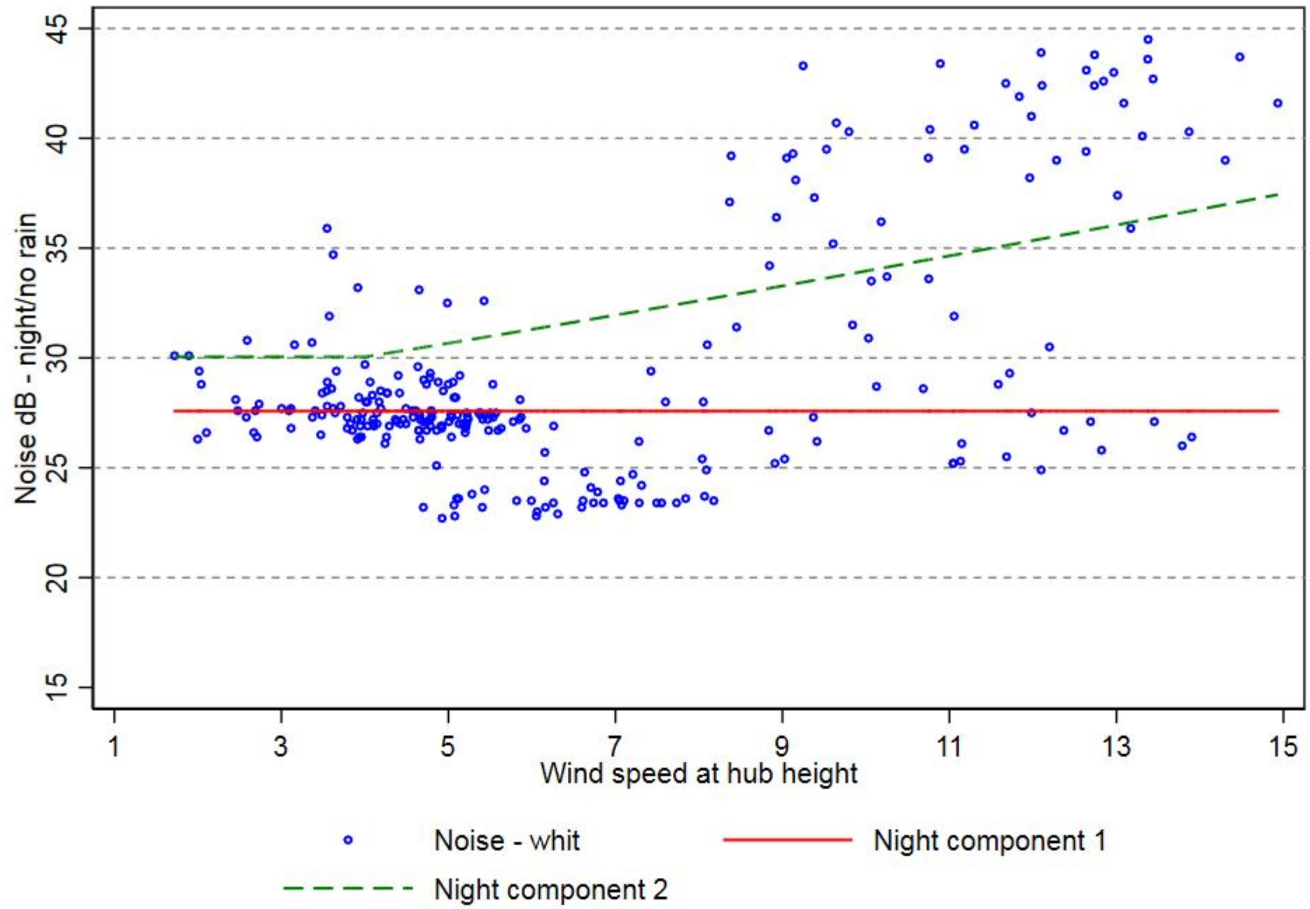
Model comparisons using BIC – location 2

Site	Components	BIC for model:			
		Beta	Normal	Lognormal	Weibull
6	1	2,180	2,156	2,078	2,314
6	2	1,911	1,901	1,872	1,997
6	2 R	1,905	1,895	1,881	1,985
6	3	1,908	1,900	1,879	1,902
6	3 R	1,889	1,892	1,871	1,946
6	4	1,890	1,886	1,881	1,862
7	1	3,988	3,897	3,734	3,950
7	2	3,576	3,558	3,516	3,607
7	2 R	3,571	3,555	3,512	3,594
7	3	3,518	3,517	3,521	3,559
7	3 R	3,522	3,521	3,495	
7	4	3,522	3,520	3,569	3,560
8	1	2,300	2,456	2,284	2,523
8	2	2,057	2,056	2,074	2,152
8	2 R	2,052	2,108	2,068	2,151
8	3	2,029	2,023	2,019	2,056
8	3 R	1,999	2,014	2,014	2,054
8	4	2,178			
9	1	4,208	4,487	4,187	4,829
9	2	3,538		3,545	3,770
9	2 R	3,533			
9	3	3,365	3,412	3,403	
9	3 R	3,354			
9	4	3,342	3,370	3,356	
10	1	2,190	2,252	2,081	2,510
10	2	1,912	1,924	1,805	2,086
10	2 R	1,908	1,921	1,827	2,081
10	3	1,669	1,694	1,663	1,807
10	3 R	1,687	1,742	1,711	1,813
10	4	1,734	1,730	1,717	1,788

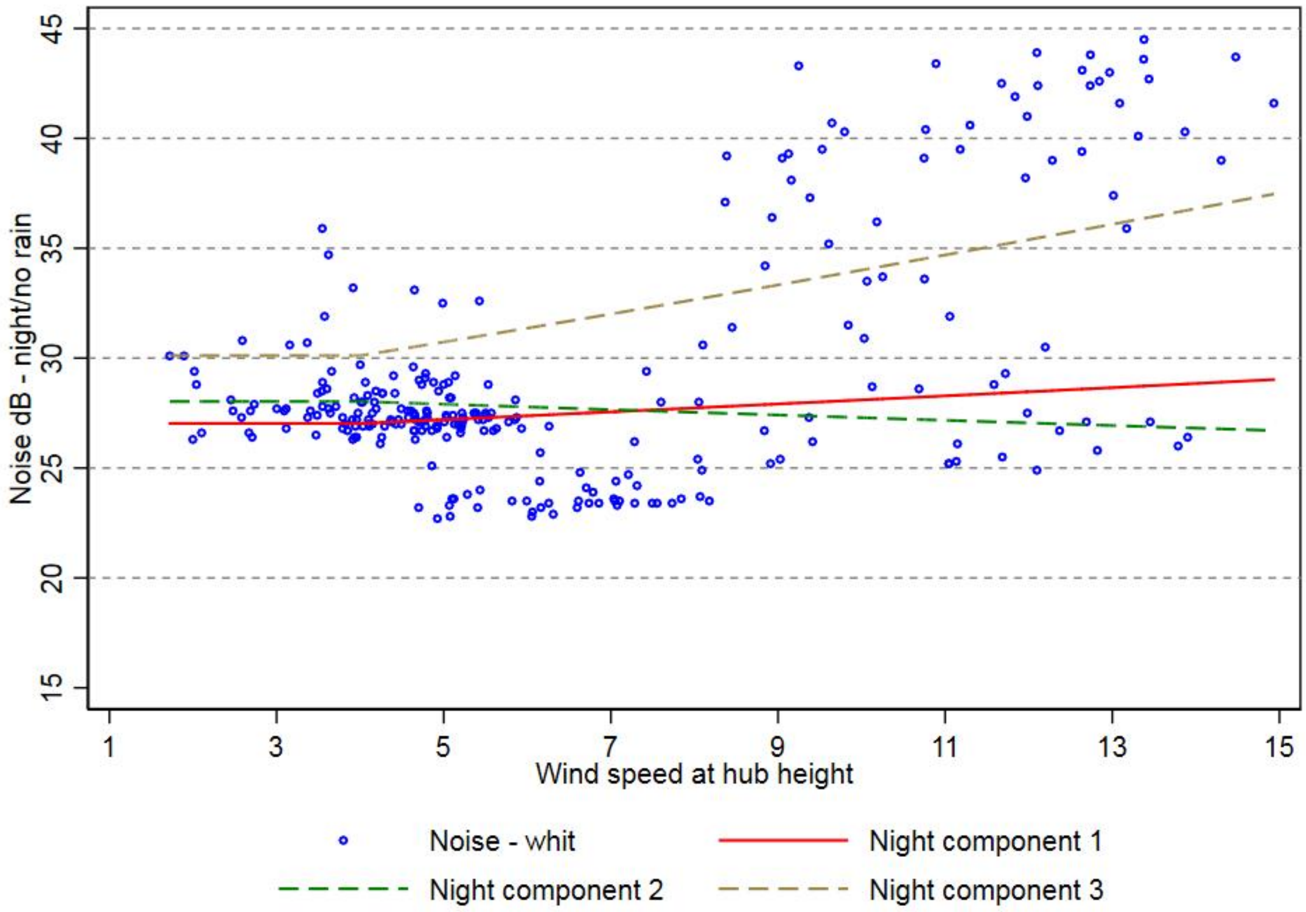
2 beta noise components (unrestricted)



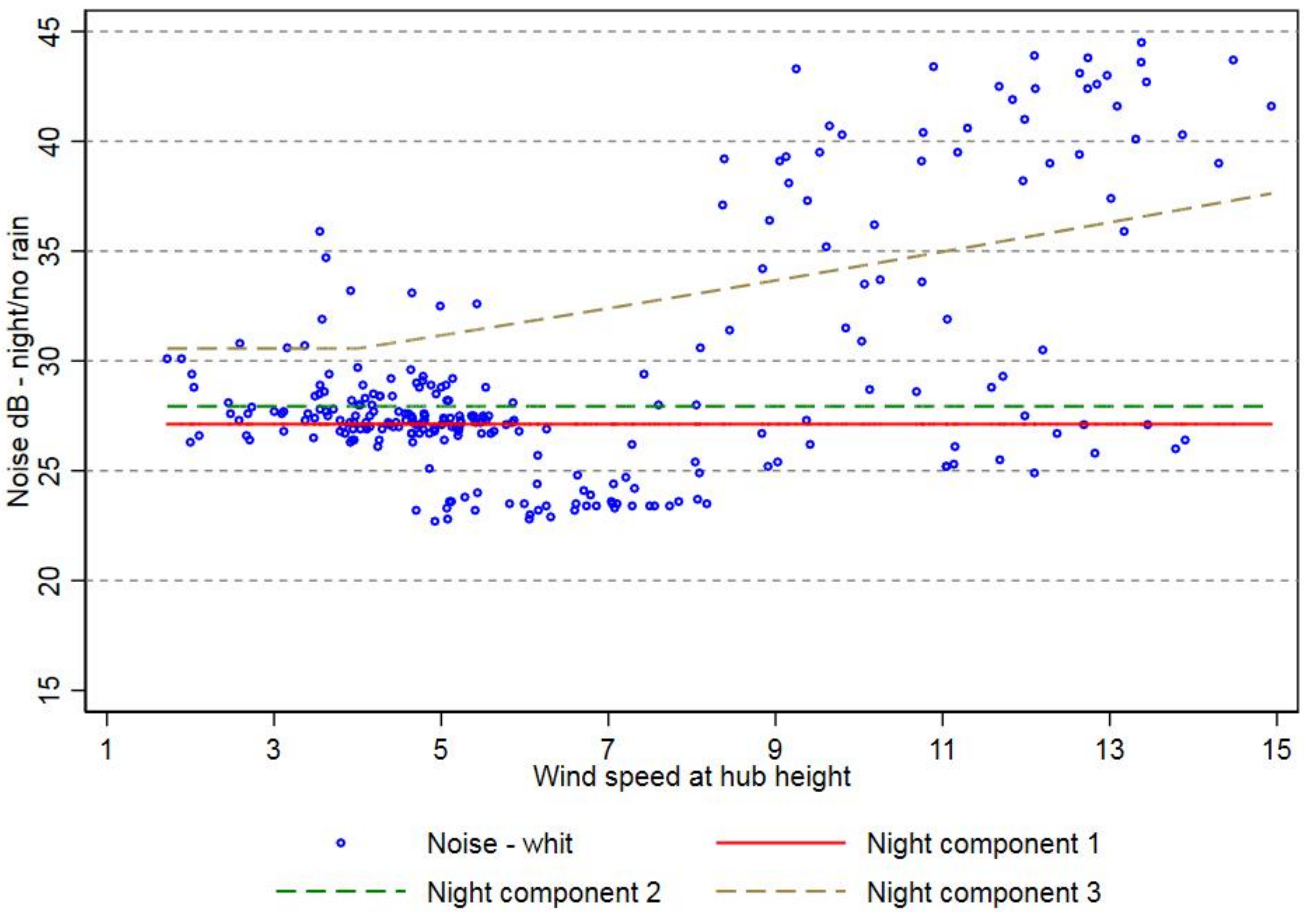
2 beta noise components (restricted)



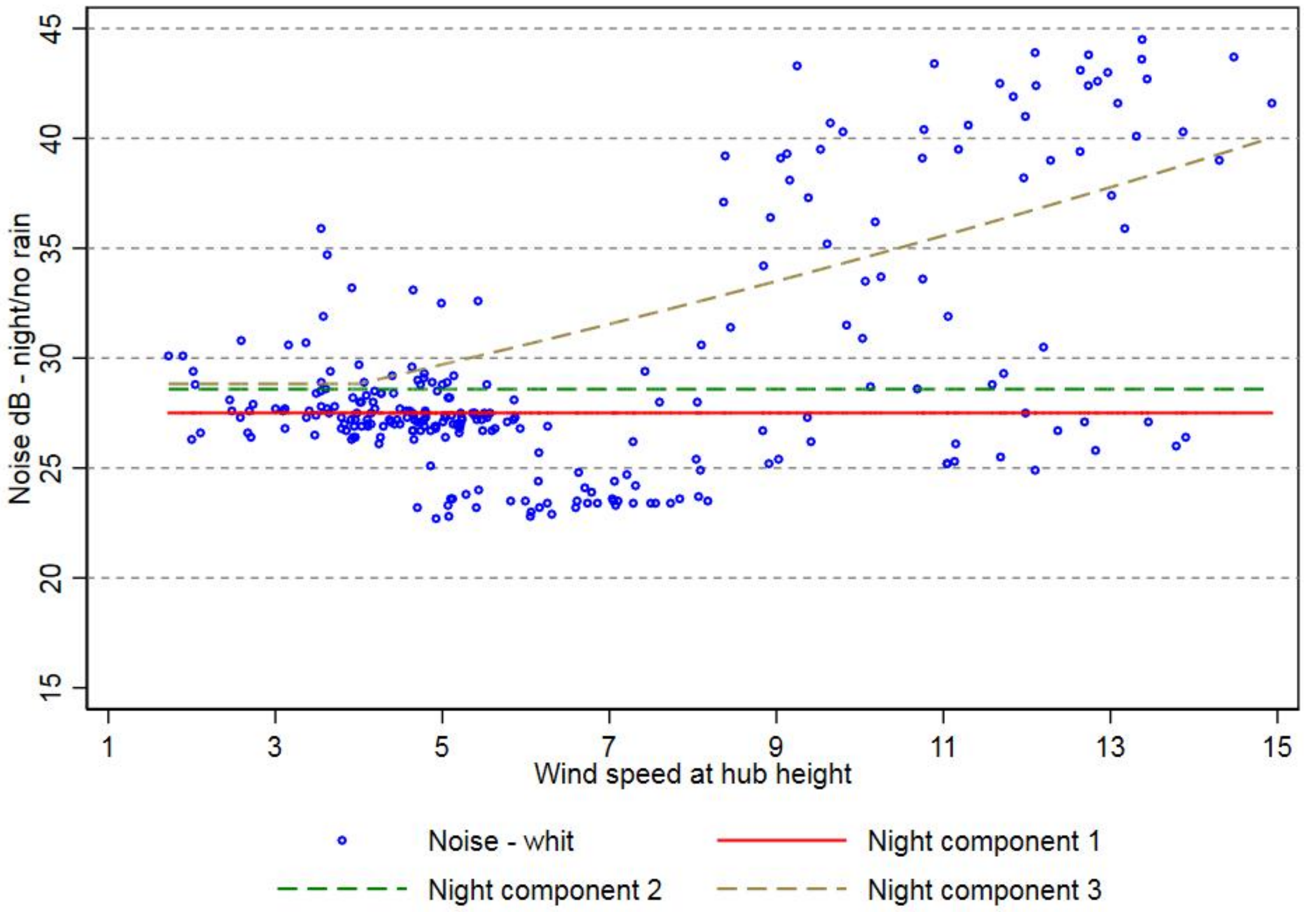
3 beta noise components (unrestricted)



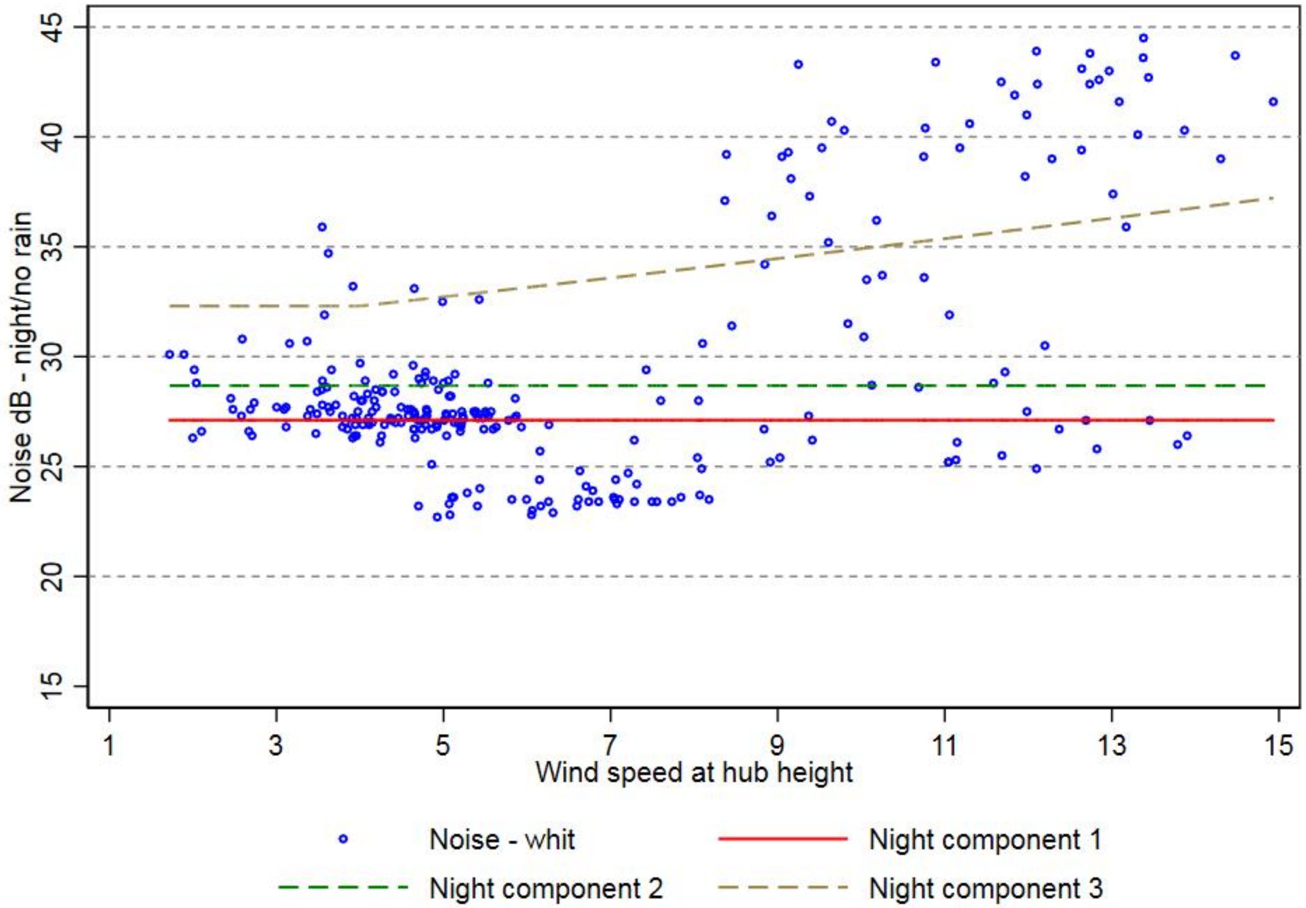
3 beta noise components (restricted)



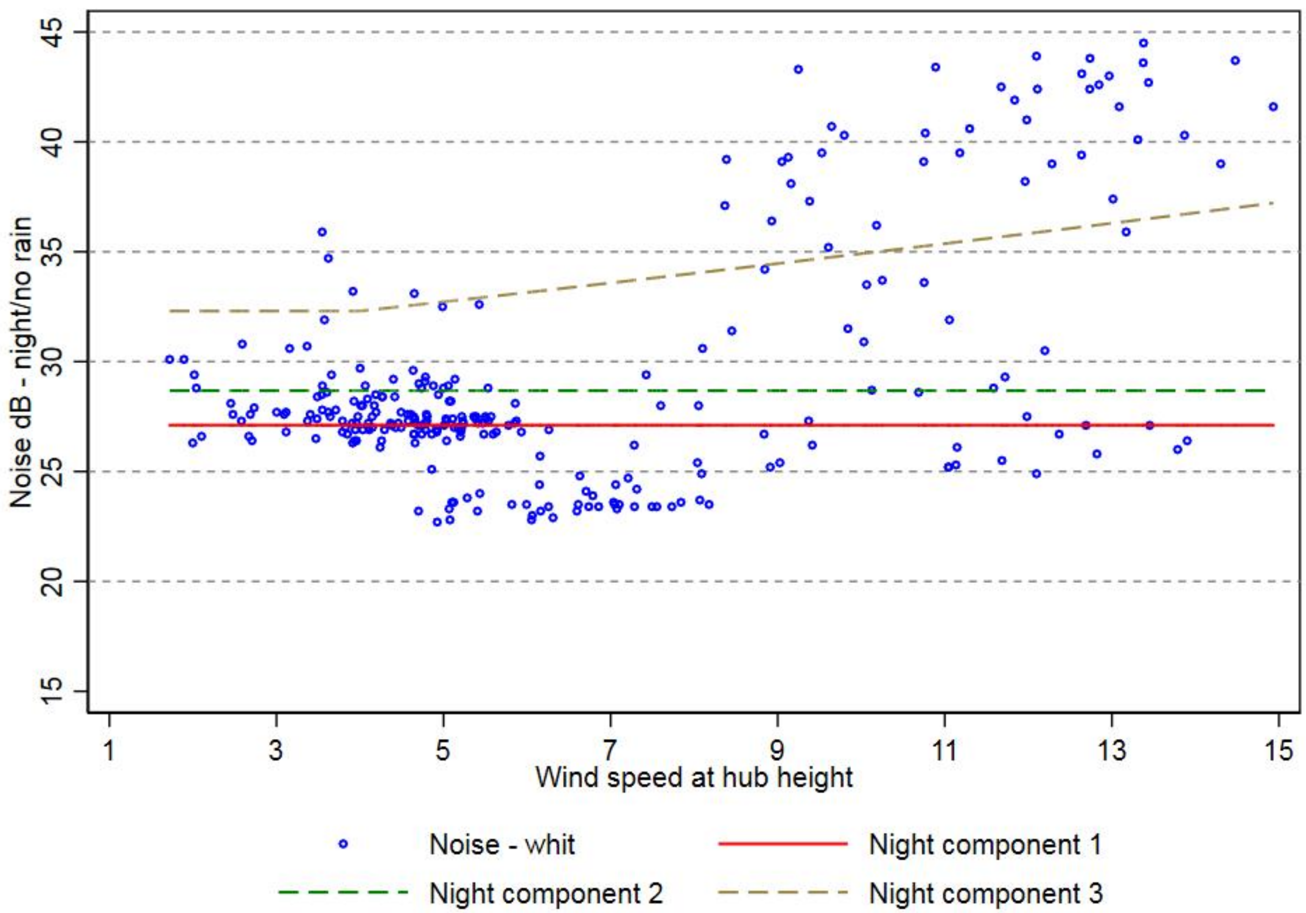
3 lognormal noise components (restricted)



3 Weibull noise components (restricted)



3 Weibull noise components (restricted)



Background noise estimates – site 1

Wind speed (m/s)	2 component specification				3 component specification	
	Unrestricted		Restricted		Restricted	
	Estimate	SE	Estimate	SE	Estimate	SE
Measure A - Background noise as base noise component (dB)						
4	27.7	0.1	27.6	0.1	27.1	0.1
5	27.6	0.1	27.6	0.1	27.1	0.1
6	27.5	0.1	27.6	0.1	27.1	0.1
7	27.4	0.2	27.6	0.1	27.1	0.1
8	27.3	0.2	27.6	0.1	27.1	0.1
9	27.2	0.3	27.6	0.1	27.1	0.1
10	27.2	0.4	27.6	0.1	27.1	0.1
11	27.1	0.5	27.6	0.1	27.1	0.1
12	27.0	0.5	27.6	0.1	27.1	0.1
Measure B - Background noise as weighted average of noise components (dB)						
4	27.9	0.1	27.8	0.1	27.8	0.1
5	28.0	0.1	28.0	0.1	28.0	0.1
6	28.5	0.1	28.5	0.1	28.4	0.1
7	29.2	0.1	29.2	0.1	29.1	0.1
8	30.2	0.1	30.2	0.1	30.1	0.2
9	31.4	0.2	31.4	0.2	31.3	0.2
10	32.6	0.2	32.6	0.2	32.6	0.2
11	33.7	0.2	33.7	0.2	33.8	0.2
12	34.7	0.3	34.8	0.3	34.9	0.3

Background noise estimates – site 2

Wind speed (m/s)	2 component specification				3 component specification	
	Unrestricted		Restricted		Restricted	
	Estimate	SE	Estimate	SE	Estimate	SE
Measure A - Background noise as base noise component (dB)						
4	23.0	0.1	23.1	0.1	23.4	0.1
5	23.3	0.2	23.1	0.1	23.4	0.1
6	23.6	0.4	23.1	0.1	23.4	0.1
7	23.9	0.5	23.1	0.1	23.4	0.1
8	24.3	0.6	23.1	0.1	23.4	0.1
9	24.6	0.8	23.1	0.1	23.4	0.1
10	25.0	1.0	23.1	0.1	23.4	0.1
11	25.3	1.1	23.1	0.1	23.4	0.1
12	25.7	1.3	23.1	0.1	23.4	0.1
Measure B - Background noise as weighted average of noise components (dB)						
4	23.6	0.1	23.7	0.1	23.5	0.1
5	24.5	0.1	24.5	0.1	23.9	0.1
6	25.8	0.1	25.8	0.1	25.4	0.2
7	27.6	0.2	27.6	0.2	27.8	0.3
8	29.7	0.2	29.7	0.2	29.6	0.3
9	32.0	0.2	31.9	0.2	31.4	0.3
10	34.2	0.3	34.0	0.2	33.8	0.3
11	36.3	0.3	36.1	0.2	36.4	0.3
12	38.3	0.3	38.1	0.3	38.9	0.4

Mean absolute differences between beta & lognormal estimates of background noise

	2 component specification		3 comp spec
	Unrestricted	Restricted	Restricted
Measure A - Background noise as base noise component (dB)			
Site 1	0.1	0.0	0.4
Site 2	0.5	0.1	0.1
Site 3	1.2	0.2	1.3
Site 4	0.0	0.0	0.1
Site 5	0.4	0.3	0.0
Site 6	0.4	0.0	0.0
Site 7	0.8	0.0	0.0
Site 8	0.5	0.5	0.0
Site 9	8.3		
Site 10	16.8	0.1	0.0
Measure B - Background noise as weighted average of noise components (dB)			
Site 1	0.0	0.0	0.5
Site 2	0.0	0.1	0.1
Site 3	0.3	0.3	0.5
Site 4	0.0	0.0	0.1
Site 5	0.2	0.2	0.2
Site 6	0.1	0.0	0.2
Site 7	0.8	0.9	0.7
Site 8	0.6	0.6	0.6
Site 9	0.5		
Site 10	0.4	0.4	0.1

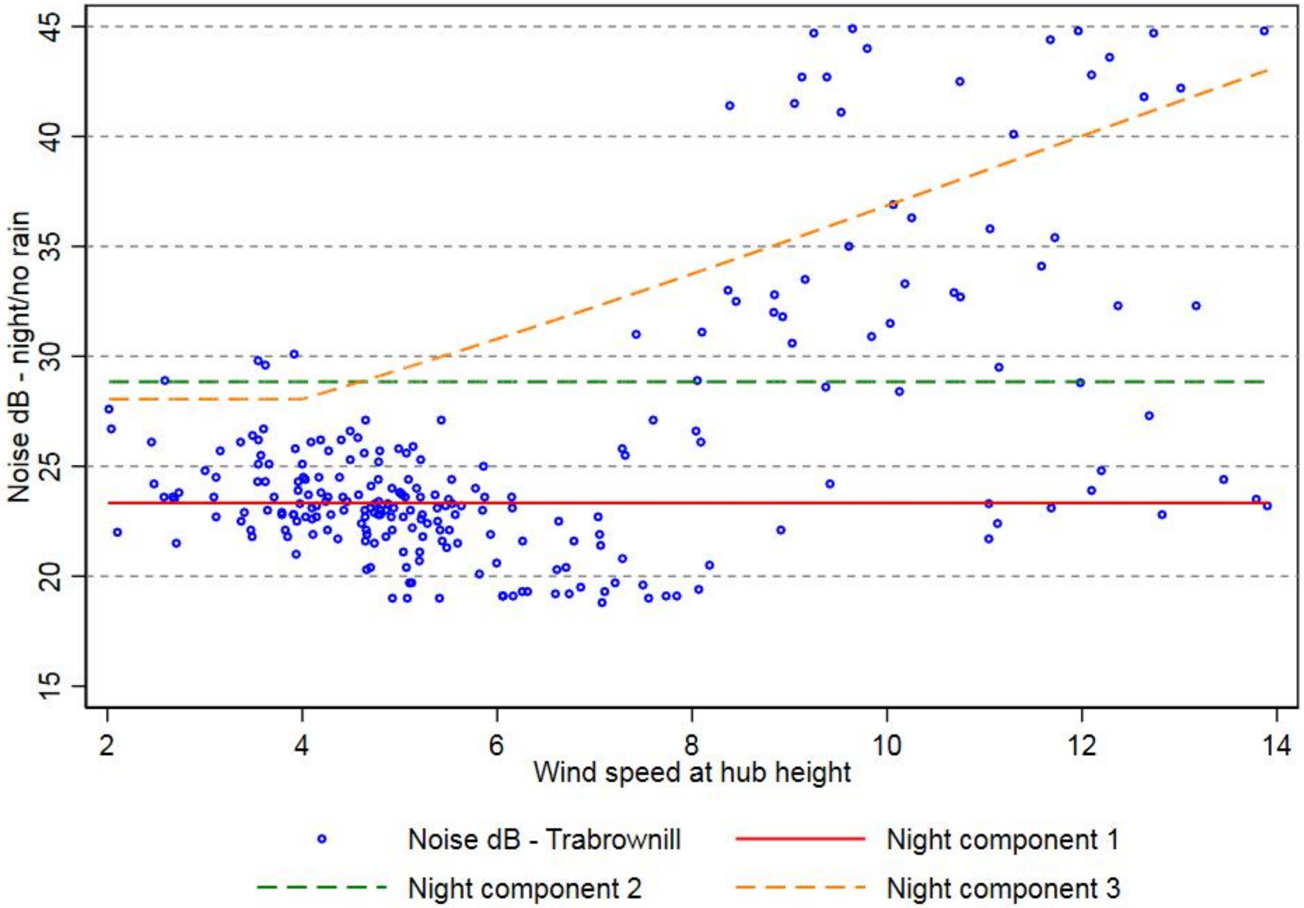
Mean absolute differences between site & panel estimates of background noise

	Beta model			Lognormal model		
	2 component specification		3 comp spec	2 component specification		3 comp spec
	Unrestricted	Restricted	Restricted	Unrestricted	Restricted	Restricted
Measure A - Background noise as base noise component (dB)						
Site 1	2.2	0.3	0.2	2.5	0.3	0.3
Site 2	0.3	0.2	0.4	0.4	0.0	0.6
Site 3	11.2	1.0	1.6	10.6	1.1	1.8
Site 4	1.3	0.4	3.9	1.6	0.2	4.2
Site 5	0.9	0.5	0.6	0.6	0.2	0.4
Site 6	4.5	0.2	0.0	5.3	0.3	0.0
Site 7	13.4	0.0	0.0	13.7	0.0	0.0
Site 8	2.6	0.2	0.1	2.3	0.3	0.0
Site 9	8.0	0.4	0.3	7.9	0.3	0.3
Site 10	19.9	0.0	0.1	19.9	0.1	0.2
Average	6.4	0.3	0.7	6.5	0.3	0.8
Measure B - Background noise as weighted average of noise components (dB)						
Site 1	0.6	0.7	0.2	0.5	0.6	0.3
Site 2	0.4	0.4	0.5	0.4	0.4	0.5
Site 3	1.6	1.1	1.1	1.7	1.3	1.2
Site 4	1.0	1.1	0.5	0.8	0.9	0.5
Site 5	0.3	0.3	0.4	0.3	0.4	0.4
Site 6	2.3	2.7	1.8	2.5	2.9	1.8
Site 7	0.8	0.7	0.5	1.2	1.1	0.4
Site 8	0.8	0.8	0.4	1.0	1.1	0.6
Site 9	1.7	1.9	0.7	2.1	2.2	0.7
Site 10	1.6	2.0	0.5	2.0	2.4	0.7
Average	1.1	1.2	0.7	1.3	1.3	0.7

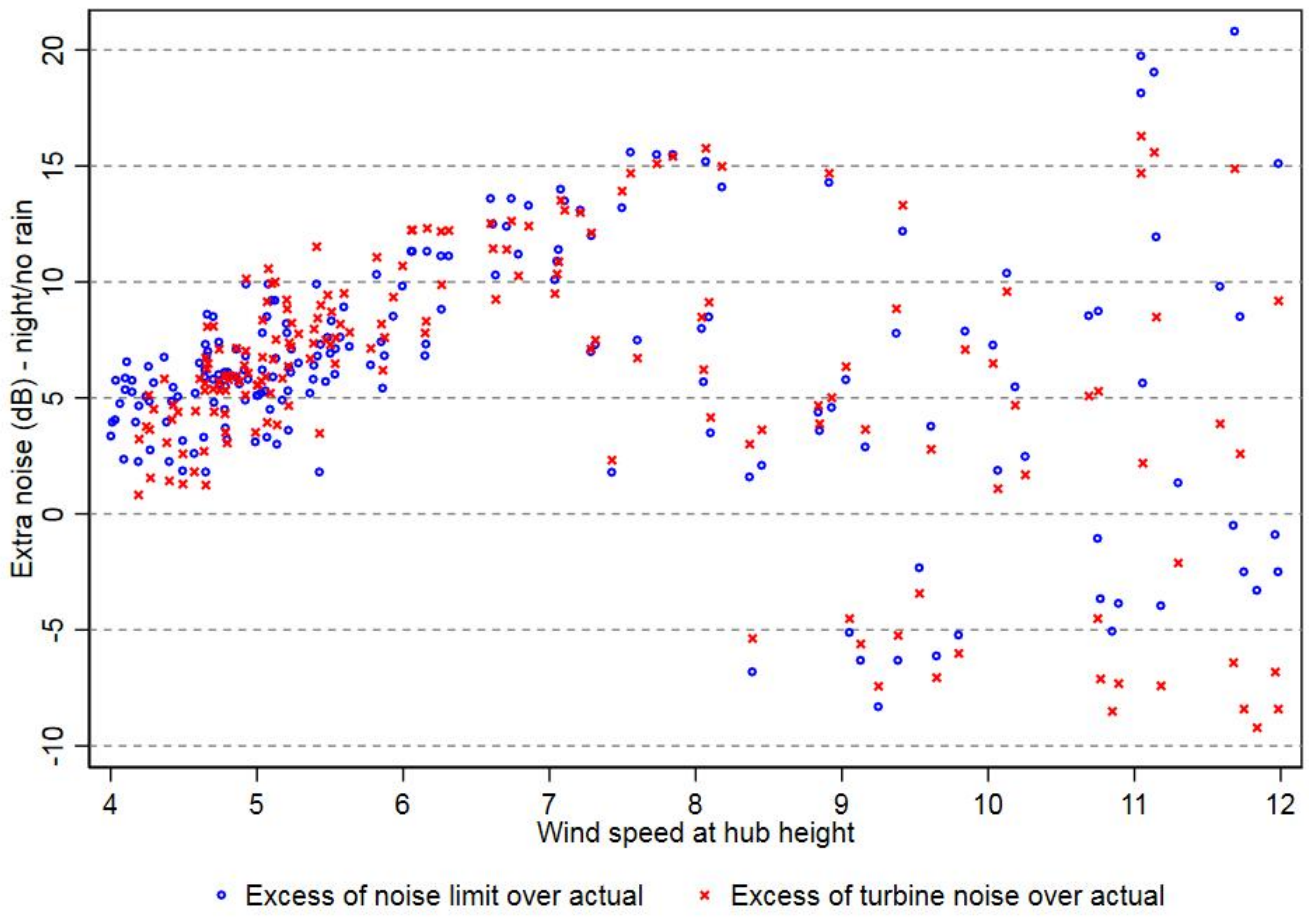
Implementing the analysis via a Stata procedure

- Prototype designed to offer a simple set of options and to generate basic graphical & tabular output for one site
- Options to control:
  - Inputs: distribution, wind speed threshold, covariates, quiet day/night, alternative noise limits, thresholds for complaints and/or significant impacts
  - Outputs: estimation results, noise components, mean probabilities of significant impacts or complaints
  - Extension to handle data presented as a spreadsheet
- Plan to add a User menu to enable users to control the procedure via drop-down choices

Estimates of noise components (3R) – site 2



Will noise limits avoid complaints – site 2?



Probability of noise complaints under alternative noise limits (10 dB threshold)

Wind speed category	Site 1	Site 2	Site 3	Site 4	Site 5
Probability of complaints for background noise limits					
Light	0.02	0.00	0.00	0.06	0.02
Moderate	0.43	0.55	0.61	0.30	0.59
Fresh	0.19	0.16	0.10	0.13	0.39
Strong	0.41	0.27	0.09	0.32	0.18
All	0.19	0.18	0.17	0.16	0.23
Probability of complaints for turbine noise limits noise limits					
Light	0.43	0.68	0.98	0.09	0.01
Moderate	1.00	0.98	0.98	0.52	0.30
Fresh	0.58	0.42	0.32	0.87	0.13
Strong	0.50	0.36	0.14	0.50	0.00
All	0.61	0.67	0.76	0.38	0.10

Lessons

- Noise assessments are carried out by acousticians with little statistical expertise and a desire for simple recipes
  - More sophisticated analysis has to be presented in a packaged form with a strong focus on critical results
- The issues are important because noise can affect a lot of people and is poorly understood
  - Need to find ways of both understanding and communicating the consequences of alternative choices about, say, noise limits
- Strong resistance to change from developers who see noise assessment as a purely mechanical exercise
- Parallels with other areas of statistical application?