MAY 10 2022

Beyond traditional wind farm noise characterisation using transfer learning

Phuc D. Nguyen; Kristy L. Hansen; Bastien Lechat; ... et. al

JASA Express Lett 2, 052801 (2022)
https://doi.org/10.1121/10.0010494

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Beyond traditional wind farm noise characterisation using transfer learning

Phuc D. Nguyen,[1,a] Kristy L. Hansen,[1] Bastien Lechat,[2] Branko Zajamsek,[1] Colin Hansen,[3] and Peter Catcheside[2]
[1]College of Science and Engineering, Flinders University, Adelaide, South Australia 5042, Australia
[2]Adelaide Institute for Sleep Health, Flinders University, Adelaide, South Australia 5042, Australia
[3]School of Mechanical Engineering, University of Adelaide, Adelaide, South Australia 5005, Australia
ducphuc.nguyen@flinders.edu.au, kristy.hansen@flinders.edu.au, bastien.lechat@flinders.edu.au, branko.zajamsek@flinders.edu.au, colin.hansen@adelaide.edu.au, peter.catcheside@flinders.edu.au

Abstract: This study proposes an approach for the characterisation and assessment of wind farm noise (WFN), which is based on extraction of acoustic features between 125 and 7500 Hz from a pretrained deep learning model (referred to as deep acoustic features). Using data measured at a variety of locations, this study shows that deep acoustic features can be linked to meaningful characteristics of the noise. This study finds that deep acoustic features can reveal an improved spatial and temporal representation of WFN compared to what is revealed using traditional spectral analysis and overall noise descriptors. These results showed that this approach is promising, and thus it could provide the basis for an improved framework for WFN assessment in the future.

[Editor: Tracianne B. Neilsen] https://doi.org/10.1121/10.0010494

Received: 6 February 2022 **Accepted:** 22 April 2022 **Published Online:** 10 May 2022

1. Introduction

The global wind industry has undergone rapid expansion in power generation capacity over the past ten years, reaching to over 22 000 wind farms and 1110 GW in 2021.[1] This fast growth is expected to continue, along with ongoing concerns regarding social,[2,3] ecological,[4,5] and environmental impacts, such as noise generated by wind turbines.[6,7] Multiple guidelines and standards[8–10] have been developed to help mitigate the effects of wind farm noise (WFN) on surrounding communities. Although these guidelines and standards have been updated regularly, the potential impact of WFN is still based on common traditional noise metrics such as A-weighted (LA) or C-weighted (LC) sound pressure levels (SPLs).[11] These aggregate metrics are clearly important indicators related to the human perception of noise. However, there remains no consensus agreement or firm evidence to support which metrics are most strongly related to human impacts and are thus most suitable for WFN assessment.[12] Moreover, prominent characteristics of WFN, such as amplitude modulation (AM), also appear to importantly contribute to annoyance[13–15] and possible sleep disturbance.[16] Consequently, it is unlikely that any single simple noise metric can adequately encapsulate both physical and psychological aspects of WFN impacts on humans, and more comprehensive and evidence-based approaches remain needed. This problem is not unique to WFN research but has also been identified as an issue in other research areas, such as sonic boom research, for which there is no internationally agreed-upon standard noise metric.[17]

Recent advances in deep learning in acoustics[18] hold significant promise for improving WFN noise assessment. In particular, Sethi *et al.*[19] recently used a pretrained deep convolutional neural network (CNN), namely, VGGish,[20] to extract feature patterns in spectrograms. Hereafter, these features are referred to as deep acoustic features. The authors showed that deep acoustic features can be used to accurately quantify variations in the natural environment across locations and time. Using a pretrained deep model (DEEP-Hybrid DataCloud project), Clar *et al.*[21] were able to characterise details of diurnal and spatial community-generated sound and noise sources. Other useful applications of machine learning models are presented in many aspects of environmental acoustics, such as outdoor sound propagation[22,23] and sound emergence of wind turbine noise (the difference in LAs between wind farm operating and non-operating condition).[24]

The purpose of this study was to investigate whether deep acoustic features can be used as an alternative to traditional acoustic features for WFN characterisation. After introducing the methods and the data sets (Sec. 2), we first sought to understand the degree of correlation between deep acoustic features and traditional acoustic features, such as spectral shape and commonly derived noise indicators (Sec. 3.2). We then explored the ability of deep acoustic features to reveal the variations in WFN characteristics across locations and time (Sec. 3.3).

[a]Author to whom correspondence should be addressed.

2. Methods

2.1 Data sets

All data sets used in this study were collected from four locations in the mid-north region of South Australia (see supplementary material[26] Table S1 and S2 for details). The first data set (data set 1) was a benchmark data set,[25] which contains 6000 10-s audio files of WFN with 40% of audio samples containing AM. These data were extracted from a data bank containing continuous data measured over 1 year at locations 1 and 2 (supplementary material[26] Table S2). In this study, data set 1 was used to evaluate the accuracy of AM detection algorithms trained with deep acoustic features. The second data set (data set 2) included all samples in data set 1 (i.e., samples extracted from locations 1 and 2) and other samples extracted from data sets measured at locations 3 and 4. Data set 2 thus contained noise samples measured at four different locations near three wind farms (see supplementary material[26] Table S2). Data set 2 contained data measured near three wind farms and four residences located between approximately 1 and 9 km from the wind farms. Data set 2 was used to investigate whether deep acoustic features can reveal the spatial fluctuations of WFN. The third data set (data set 3) was extracted from data measured at location 2 over 1 year. To reduce the computational time, we extracted 10-s audio files from 10-min samples. In total, the third data set contained over 50 000 10-s audio samples. Data set 3 was used to evaluate whether deep acoustic features can reveal the temporal fluctuations of WFN. More details regarding wind farm characteristics and measurement locations were provided in supplementary Tables S1 and S2.[26] Other information regarding equipment and measurement setup can be found in our previously published work.[27] Regarding data cleaning, although we removed all data containing rain contamination and farming machine noise, we were unable to separate WFN from ambient noise. This is a current issue in the WFN research area, and there is no current validated method to address this problem.[12] A recent study conducted by Gloaguen *et al.*[24] showed promising results to address this challenge, and this approach could improve WFN characterisation in future studies.

2.2 VGGish

We used the VGGish model, which is a CNN model that has been pretrained on more than 2×10^6 YouTube videos to predict over 600 audio event classes.[20] The architecture of the VGGish model is shown in Fig. 1(a). The VGGish model includes a single channel input layer, followed by four convolutional (CONV) layers and three fully connected layers. VGGish is a variant of the well-known VGG model, in particular configuration A with 11 weight layers. To prepare the input for the VGGish model, the audio signal was framed into zero overlapping windows of 0.96 s. Each window was converted into a spectrogram using a short-time Fourier transform with a window size of 25 ms, a hop length of 10 ms, and a periodic Hanning window. A mel spectrogram was then computed by mapping the spectrogram to 64 mel bins covering the range of 125–7500 Hz. To avoid calculating a logarithm of zero, a log mel spectrogram was computed by adding a log offset value of 0.001. The result of each window was a two-dimensional (2D) spectrogram image of 96×64 pixels (i.e., 96 frames by 64 mel bands). This image was used as the input to the VGGish model. For other details regarding the VGGish model, we refer the reader to the source code provided by the TensorFlow team.[28]

A pretrained model was used to save computational and human resources involved with the training and validation of a new model. Although the VGGish model was not trained to identify WFN indicators specifically, it was trained using audio files that contain unique features of noise, such as swoosh, swish, and thump, widely associated with WFN.[29] Application of this pretrained model thus could be particularly useful when applied to characterise WFN. Another benefit of using the pre-trained VGGish model is that it is unlikely to be overfitted because it has been trained using a massive data set and is thus capable of classifying a wide range of noise types not necessarily represented within our potentially somewhat location-specific data set that may differ compared to other regions. In other words, this model has the potential to perform well when applied to different data sets containing WFN with a wider range of noise characteristics. Also, the VGGish model was trained using mel spectrogram images, which better reflect how humans perceive sound compared to conventional spectrograms calculated using a short-time Fourier transform. The mel spectrogram represents a psychoacoustic scale, and thus the resulting deep acoustic features are expected to show stronger relationships with human responses (i.e., annoyance) compared to conventional spectrograms, although further investigation is needed to verify this. The limitation of the VGGish model is that the upper bound frequency of the input mel spectrogram was high for WFN, which typically contains frequencies below 2000 Hz. Removing these frequencies could thus potentially improve model performance. Also, the frequency content of WFN in the infrasonic and very low frequency ranges (i.e., below 125 Hz) is removed when using the VGGish model. However, this could still be acceptable because wind farm infrasound is normally well below the human hearing threshold,[30] and no studies to date have shown that it has adverse effects on humans. Also, the most commonly identified problematic component of WFN is AM, which often occurs within the frequency range between 200 and 800 Hz.[31] This information is thus well captured by the VGGish model.

2.3 Dimensionality reduction methods

We used principal component analysis (PCA)[32] to reduce from higher- to lower-dimensional acoustic features. This process can efficiently remove highly correlated acoustic features that are redundant, resulting in an improvement in the

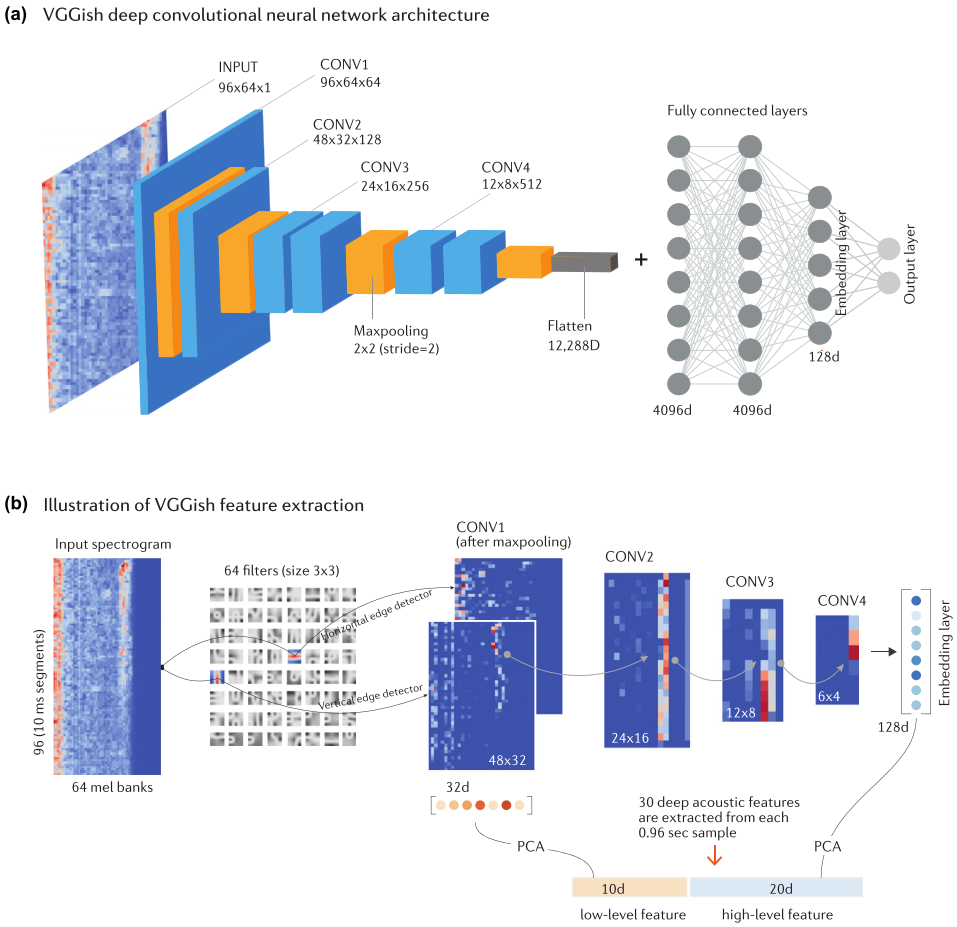

Fig. 1. VGGish and deep acoustic features. (a) The architecture of the VGGish model. (b) An illustration of the VGGish feature extraction. Curved paths represent how the input spectrogram is convoluted with the filters to create new images in the CONV layers. We finally combine the features in the CONV1 and embedding layers to represent the acoustic information of the input noise sample. These combined features are referred to as deep acoustic features.

machine learning algorithm performance and reducing the risk of overfitting. In our analysis, the number of features extracted from the CONV1 and embedding layers was reduced from 32 to 10 and from 128 to 20, respectively [Fig. 1(b)]. The number of principal components was chosen such that most (95%) of the variance within the high-dimensional data was preserved in the lower-dimensional data (i.e., the cumulative explained variance was above 0.95 in this study). Thus, the total number of combined low-level and high-level deep acoustic features was 30 [Fig. 1(b)]. To visualise deep acoustic features on a two-dimensional plot, we used the uniform manifold approximation and projection (UMAP)[33] to further reduce dimensionality from 30 deep acoustic features to two features. The benefit of UMAP compared to PCA is that it can preserve both global and local geometry of the data, so it is expected to provide superior visualisation of WFN components. The benefit of using UMAP compared to other projection methods, such as t-distributed stochastic neighbour embedding (t-SNE), is that UMAP is more computationally efficient and better preserves the global structure of the data compared to t-SNE (see supplementary Fig. S1 for details).[26]

2.4 Performance of deep acoustic features

To evaluate the ability of deep acoustic features to identify WFN AM, we used a recent advanced and successful machine learning method called extreme gradient boosting (XGBoost)[34] to identify AM in data set 1. The XGBoost algorithm uses multiple decision trees to capture nonlinear relationships between input variables (acoustic features) and output (AM vs no AM) to make an ensemble prediction. We also used XGboost classifiers to evaluate the ability of deep acoustic features to reveal spatial and temporal dependencies of WFN. All XGboost classifiers were trained on 80% and evaluated on 20% of the data. Hyperparameters of XGboost classifiers, such as *learning_rate* and *max_depth*, were optimised with the Bayesian hyperparameter tuning method.[35] The searching space and optimised hyperparameters of the AM classifiers are

provided in supplementary Tables S3 and S4.[26] We reported the performance of the classifiers on the test sets throughout this study.

2.5 Statistical analysis

Statistical analyses, including Pearson's correlations and receiver operating characteristic (ROC) curve analyses, were performed using R version 4.0.0.[36] ROC area under curve (AUC) and F1-score were used as the main classifier performance metrics throughout this study. All visualisation was implemented in R using *ggplot2* and *circlize* packages. All grouped data are reported as mean ± SD.

2.6 Data and code availability

The data sets used in this study are available in Ref. 37. The source code used to extract deep acoustic features and generate the main figures is published in Ref. 38.

3. Results

3.1 Deep acoustic features

We proposed a different approach from previous studies,[19] in which features in CONV and embedding layers were extracted instead of features in the embedding layer only. We hereafter refer to features extracted from CONV layers as low-level acoustic features and features extracted from the embedding layer as high-level acoustic features [Fig. 1(b)]. While high-level acoustic features can capture global acoustic content of an input spectrogram, such as energy distribution, as a function of frequency and time, we also expected that low-level acoustic features could indicate local patterns in the input spectrogram, such as intermittent acoustic features represented by vertical and horizontal lines. Combining these features was expected to capture both the general and detailed characteristics of the noise. To investigate the meaning of each layer in the deep model and to test whether low-level acoustic features indeed capture detailed patterns of the input spectrogram, we show a case study in Fig. 1(b). In this case study, a spectrogram of a 0.96-s audio recording was input to the deep model. The input spectrogram was convoluted with 64 filters to create 64 new images in the CONV1 layer. Each filter was trained to detect different patterns of the input spectrogram. To visualise this, we highlighted two fundamental filters and showed the corresponding images in the CONV1 layer [Fig. 1(b)]. The filters successfully detected vertical (distribution of acoustic energy over frequency bands) and horizontal (intermittent acoustic energy over time) patterns of the spectrogram. There were 64 filters in the CONV1 layer that can thus detect other detailed patterns of the input spectrogram. These new spectrogram images representing local patterns of the input were then passed to deeper layers (i.e., CONV2–4 layers), where the number of filters were doubled after each layer, resulting in much more detailed and comprehensive patterns of the input spectrogram. We reasoned that this combination of low- and high-level acoustic features [Fig. 1(b)] can comprehensively capture both detailed and general information about the input noise sample. However, features extracted from deeper CONV layers were very localised rather than global. In this study, we only used the low-level acoustic features extracted from the CONV1 layer [Fig. 1(b)]. The explanations for this are analysed in detail in Sec. 3.2. Finally, the deep acoustic features include 10 low-level features and 20 high-level features.

3.2 Deep acoustic features reveal noise characteristics

To investigate potential relationships between deep and traditional acoustic features and WFN unique features (i.e., AM characteristics), we ran a pairwise correlation analysis, in which all possible pairs between these features were determined, and their corresponding Pearson's correlation coefficients, r, were then calculated. The traditional acoustic features included spectral shape features (i.e., spectral slope, spectral spread, spectral flux) and environmental noise indicators (i.e., LA and LC). The AM features included AM fundamental frequency and AM depth. A full list of these features and their physical meaning is provided in supplementary Table S5.[26] The pairwise correlation coefficients were presented in a chord diagram[39] as shown in Fig. 2(a). We observed moderate to strong correlations between these features. Both low- and high-level acoustic features were correlated with traditional acoustic features and AM features, indicating that both low- and high-level acoustic features are useful for capturing information about the noise character. Deep acoustic features were correlated with the most important AM features, such as the AM fundamental frequency, *peakloc*; AM strength (depth), *DAM*, *AMfactor*; rising and decay slopes of AM peaks, *pos_slope*, *neg_slope*; and variations of unweighted SPL in octave bands centred at 1000 Hz, *L*_1000. Furthermore, deep acoustic features were also correlated with common traditional acoustic features, such as LAs, LC/LA ratio, G-weighted SPL (LG)/LA, and spectral shape features, such as spectral spread and spectral slope. This is striking as the correlation analysis reveals that the deep acoustic features are not only able to represent traditional acoustic features but are also related to unique characteristics of WFN.

To further investigate the ability of deep acoustic features to detect AM, we trained an XGBoost classifier to detect AM in data set 1 and measured the performance using AUC (see Sec. 2). We also compared the performance of deep acoustic features and features extracted from CONV and embedding layers. The models trained with features extracted from CONV1 and embedding layers had higher performance compared with those trained with features

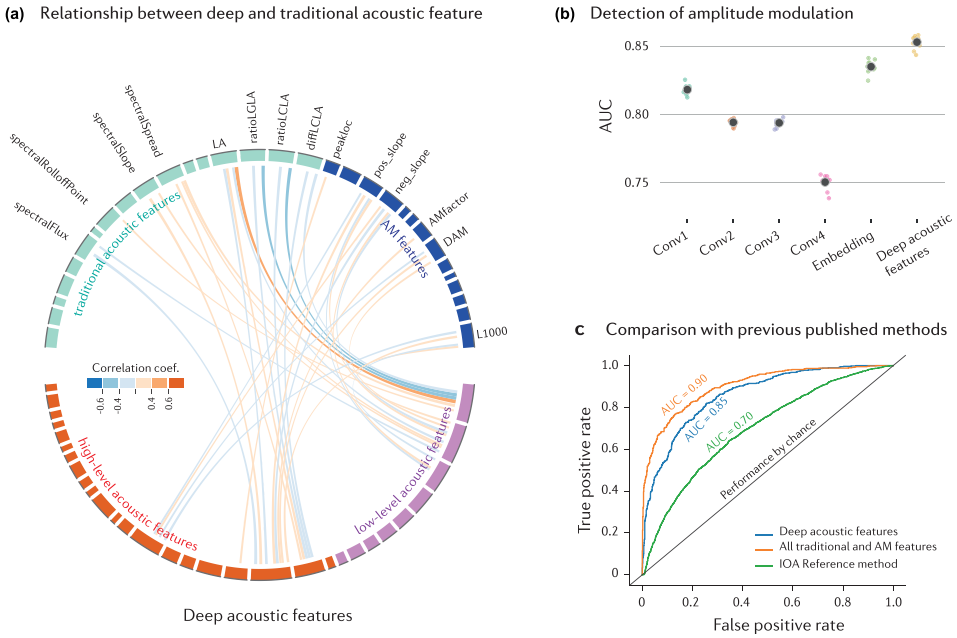

Fig. 2. Deep acoustic features reveal unique noise characteristics. (a) Relationship between deep acoustic features and traditional acoustic and AM features. The colours of the curves connected between the two features indicate the degree of correlation. (b) Performance of models trained with features extracted from CONV1–4 layers, embedding layers, and deep acoustic features. (c) Performance of the model trained with deep acoustic features is compared with previously published methods.

extracted from CONV2-4 layers [Fig. 2(b)]. The deep acoustic features showed the best performance. This was expected, as the deep acoustic features were a combination of features extracted from CONV1 and embedding layers [Fig. 1(b)].

Finally, we compared the performance of the deep acoustic feature method with previously published methods for identifying AM, including the reference method developed by the Institute of Acoustics UK[31] and our previously published method[25] using all traditional acoustic and AM features. The AUCs of the deep acoustic feature method for the training and testing data sets were 0.86 and 0.85 (F1-score = 0.71), respectively. The AUCs of the model trained with traditional acoustic and AM features for the training and testing data sets were both approximately 0.9 (F1-score = 0.76). The accuracy of the model trained with deep acoustic features (AUC = 0.85, F1-score = 0.71) was higher than that of the reference method (AUC = 0.7, F1-score = 0.59) [Fig. 2(c)]. However, the deep acoustic feature model had a lower accuracy than the model trained with traditional acoustic and AM features. This was expected, as these features were designed specifically for AM detection, including features from other AM detectors, such as *DAM*, *AMfactor*, and prominence ratio. Furthermore, all of these features were carefully selected by acoustic experts for the AM detection task. It is thus not surprising that its performance was the best, but the main advantage of deep acoustic features is that it had high performance without involving an acoustic expert. We expected that deep acoustic features could be used to detect other unique characteristics of WFN, such as tonal and impulsive characteristics.

3.3 Deep acoustic features reveal spatial and temporal dependencies of WFN

We further explored whether deep acoustic features can be used to identify spatial dependencies of WFN (e.g., linked to the geometry of the wind farm and/or the topography of the site) as well as temporal dependencies (e.g., linked to the weather conditions and/or the noise production of the wind turbines). Figure 3(b) focuses on spatial dependencies and presents four clear clusters corresponding with noise measured at four locations, as shown 1–4 in Fig. 3(a). The distances between these clusters could be used to assess the degree by which the noise measurements differ between locations. To evaluate the performance of deep acoustic features, we trained an XGboost model to classify noise at different locations. We found that the performance of the deep acoustic feature model was remarkable, with an AUC = 0.98±0.003 (F1-score = 0.98 ± 0.002). The performance of the model based on traditional and AM features had an AUC = 0.76 ± 0.008 (F1-score = 0.6 ± 0.01) and lower AUC = 0.71±0.01 (F1-score = 0.51 ± 0.01) if only LA and LC were used. These results indicate that the traditional acoustic features used here are not as sensitive to spatial variations in WFN as deep acoustic features. Although we observed differences in noise characteristics measured at different locations using the deep acoustic features, the underlying reason for these differences remains unclear, and the results could be affected by local ambient noise, distance to the wind farm, and the number of wind turbines. Larger and more suitable data sets are needed to investigate whether the deep acoustic features are sensitive to these changing conditions.

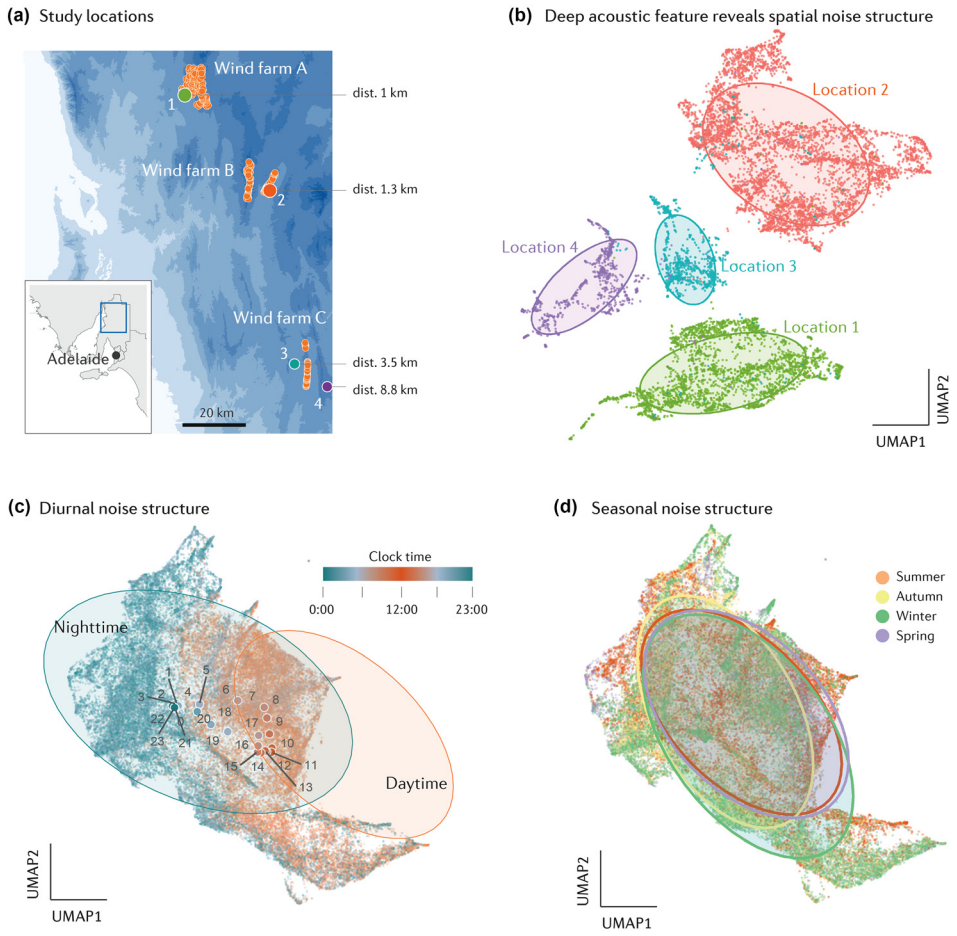

Fig. 3. Spatial and temporal characteristics of WFN. (a) Noise measured at four locations in South Australia. (b) The numbers of data points at locations 1–4 are 3000 (N1 and N2) and 1000 (N3 and N4). The elliptical shape shows one standard deviation (SD) of the data in the cluster. (c) Colours of data points indicate clock time. The numbers indicate the centroids of data points at different clock times from 0 to 23. The number of data points is 56 356. (d) Seasonal characteristics of the noise with summer (December–February), autumn (March–May), winter (June–August), and spring (September–November).

Daytime and nighttime noise dependencies were expected to be different due to atmospheric stability, which changes between nighttime and daytime. This was revealed in Fig. 3(c) using deep acoustic features. The diurnal pattern corresponds to the UMAP1 axis, where the nighttime data points were clustered on the left, while the daytime data points were clustered on the right. Interestingly, we also observed a transition between the daytime and nighttime noise dependencies. For example, the centroid of data points at clock times close to midnight (22:00 to 2:00) and midday (10:00 to 15:00) were well separated into two groups [Fig. 3(c)]. The transition hours were distributed between these two groups. These findings indicate that the deep acoustic features are very sensitive to small changes in the diurnal noise dependencies. Although the dependencies of the noise at different seasons were also expected to be different, due to changes in weather conditions, this was, however, not observed [Fig. 3(d)] except for a slightly higher SD in winter, and this observation is likely due to specific seasonal effects. It is possible that the weather conditions were not significantly different across seasons at our measurement locations. For example, there is no snow cover during winter, resulting in a reduced change in the ground impedance compared to other locations in colder climates. Furthermore, the average temperature difference between winter and summer during the measurement period was approximately 15 °C.[27] Finally, the wind farm power output between months was not significantly different (see supplementary Fig. S2 for details).[26] These factors could explain the small changes in noise characteristics between seasons.

4. Discussion

This work shows that deep acoustic features can produce similarly meaningful representations of noise as traditional acoustic features, such as LA, LC, LG, and spectral shape indicators, but can also represent unique spatial and temporal dependencies, such as AM. The use of deep acoustic features, as presented in this study, could provide a more

comprehensive approach not only to characterise the unique features of WFN, such as AM, but also to reveal the variability of noise characteristics as a function of time and location.

Although the current approach to assess WFN is mainly based on common metrics, such as LA or LC, there is still debate around these metrics as they do not capture unique noise characteristics associated with WFN.[12] The advantages of traditional acoustic features are that they are already widely implemented, accepted, and interpretable. However, these weighted measures have largely evolved around measurement and interpretation convenience, with uncertain relationships with human impacts. Given recent advances in computational resources, publicly available deep models, and especially the ability of deep acoustic features to provide more detailed information, more systematic and evidence-based measures beyond traditional approaches for assessing WFN should be considered. For example, the baseline characteristic of the noise can be established using deep acoustic features extracted from data measured during the pre-construction phase of wind farms. The operational characteristic of the noise during the operational phase of a wind farm can also be monitored. The difference in the characteristics of the noise during the pre-construction and operational phases of a wind farm could be estimated and used as an overall indicator to quantify how a wind farm alters the noise characteristics with potential impacts on amenity in surrounding communities. A similar approach has indeed been successfully applied in other systems, such as for civil[40] and mechanical structural health monitoring[41] systems. We thus anticipate that this approach will have significant utility for more comprehensive evaluation of environmental noise impacts.

5. Conclusion

We conclude that using deep acoustic features is a useful approach for characterising WFN and environmental noise in general, which captures relevant acoustic characteristics in frequencies between 125 and 7500 Hz. The noise characteristics presented in this study included both WFN and ambient noise. We showed that deep acoustic features represent both overall physical properties and characteristics unique to WFN. Deep acoustic features can clearly reveal the spatial and temporal characteristics of WFN, providing more detailed insight into the noise character than traditional acoustic features. Future use of deep acoustic features holds major promise for comprehensive assessment of the overall character of environmental noise in regions surrounding wind farms.

Acknowledgments

The authors gratefully acknowledge financial support from the Australian Research Council, Projects DP120102185 and DE180100022, and the National Health and Medical Research Council, Project 1113571. DPN was supported by a Flinders University Research Scholarship (FURS) for this work.

References and links

[1]M. Pierrot, "The windpower: Wind energy market intelligence," https://www.thewindpower.net/ (Last viewed December 24, 2021).
[2]M. Wolsink, "Wind power implementation: The nature of public attitudes: Equity and fairness instead of 'backyard motives,'" Renew. Sustain. Energy Rev. **11**(6), 1188–1207 (2007).
[3]S. Krohn and S. Damborg, "On public attitudes towards wind power," Renew. Energy **16**(1), 954–960 (1999).
[4]M. Thaker, A. Zambre, and H. Bhosale, "Wind farms have cascading impacts on ecosystems across trophic levels," Nat. Ecol. Evol. **2**(12), 1854–1858 (2018).
[5]E. Schuster, L. Bulling, and J. Köppel, "Consolidating the state of knowledge: A synoptical review of wind energy's wildlife effects," Environ. Manage. **56**(2), 300–331 (2015).
[6]G. Micic, B. Zajamsek, L. Lack, K. Hansen, C. Doolan, C. Hansen, A. Vakulin, N. Lovato, D. Bruck, C. L. Chai-Coetzer, J. Mercer, and P. Catcheside, "A review of the potential impacts of wind farm noise on sleep," Acoust. Aust. **46**, 87 (2018).
[7]T. Liebich, L. Lack, K. Hansen, B. Zajamšek, N. Lovato, P. Catcheside, and G. Micic, "A systematic review and meta-analysis of wind turbine noise effects on sleep using validated objective and subjective sleep assessments," J. Sleep Res. **30**, e13228 (2021).
[8]AS 4959, "Acoustics—Measurement, prediction and assessment of noise from wind turbine generators" (Standards Australia, Sydney, Australia, 2010).
[9]NZS 6808, "Acoustics—Wind farm noise" (Standards New Zealand, Wellington, New Zealand, 2010).
[10]L. S. Søndergaard, C. Thomsen, and T. H. Pedersen, "Prominent tones in wind turbine noise–round-robin test of the IEC 61400-11 and ISO/PAS 20065 methods for analysing tonality content," in *Proceedings of the 8th International Conference on Wind Turbine Noise*, Lisbon, Portugal (June 12–14, 2019).
[11]C. H. Hansen, C. J. Doolan, and K. L. Hansen, *Wind Farm Noise: Measurement, Assessment and Control*, 1st ed. (John Wiley & Sons Ltd., Chichester, UK, 2017).
[12]C. Hansen and K. Hansen, "Recent advances in wind turbine noise research," Acoustics **2**, 171–207 (2020).
[13]S. Lee, K. Kim, W. Choi, and S. Lee, "Annoyance caused by amplitude modulation of wind turbine noise," Noise Control Eng. J. **59**(1), 38–46 (2011).
[14]B. Schäffer, S. J. Schlittmeier, R. Pieren, K. Heutschi, M. Brink, R. Graf, and J. Hellbrück, "Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study," J. Acoust. Soc. Am. **139**(5), 2949–2963 (2016).
[15]C. Ioannidou, S. Santurette, and C.-H. Jeong, "Effect of modulation depth, frequency, and intermittence on wind turbine noise annoyance," J. Acoust. Soc. Am. **139**(3), 1241–1251 (2016).
[16]M. G. Smith, M. Ögren, P. Thorsson, L. Hussain-Alkhateeb, E. Pedersen, J. Forssén, J. Ageborg Morsing, and K. Persson Waye, "A laboratory study on the effects of wind turbine noise on sleep: Results of the polysomnographic witnes study," Sleep **43**, zsaa046 (2020).

[17]J. DeGolia and A. Loubeau, "A multiple-criteria decision analysis to evaluate sonic boom noise metrics," J. Acoust. Soc. Am. **141**(5), 3624 (2017).
[18]M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, and C.-A. Deledalle, "Machine learning in acoustics: Theory and applications," J. Acoust. Soc. Am. **146**(5), 3590–3628 (2019).
[19]S. S. Sethi, N. S. Jones, B. D. Fulcher, L. Picinali, D. J. Clink, H. Klinck, C. D. L. Orme, P. H. Wrege, and R. M. Ewers, "Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set," Proc. Natl. Acad. Sci. U.S.A. **117**(29), 17049–17055 (2020).
[20]S. Hershey, S. Chaudhuri, D. P. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold, M. Slaney, R. Weiss, and K. Wilson, "CNN architectures for large-scale audio classification," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA (March 5–9, 2017), pp. 131–135.
[21]S. N. Clark, A. S. Alli, R. Nathvani, A. Hughes, M. Ezzati, M. Brauer, M. B. Toledano, J. Baumgartner, J. E. Bennett, J. Nimo, J. Bedford Moses, S. Baah, S. Agyei-Mensah, G. Owusu, B. Croft, and R. E. Arku, "Space-time characterization of community noise and sound sources in Accra, Ghana," Sci. Rep. **11**(1), 11113 (2021).
[22]C. R. Hart, D. K. Wilson, C. L. Pettit, and E. T. Nykaza, "Machine-learning of long-range sound propagation through simulated atmospheric turbulence," J. Acoust. Soc. Am. **149**(6), 4384–4395 (2021).
[23]C. R. Hart, N. J. Reznicek, D. K. Wilson, C. L. Pettit, and E. T. Nykaza, "Comparisons between physics-based, engineering, and statistical learning models for outdoor sound propagation," J. Acoust. Soc. Am. **139**(5), 2640–2655 (2016).
[24]J.-R. Gloaguen, D. Ecotière, B. Gauvreau, A. Finez, A. Petit, and C. L. Bourdat, "Automatic estimation of the sound emergence of wind turbine noise with nonnegative matrix factorization," J. Acoust. Soc. Am. **150**(4), 3127–3138 (2021).
[25]P. D. Nguyen, K. L. Hansen, B. Lechat, P. Catcheside, B. Zajamsek, and C. H. Hansen, "Benchmark characterisation and automated detection of wind farm noise amplitude modulation," Appl. Acoust. **183**, 108286 (2021).
[26]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0010494 for detailed methodology.
[27]P. D. Nguyen, K. L. Hansen, P. Catcheside, C. Hansen, and B. Zajamsek, "Long-term quantification and characterisation of wind farm noise amplitude modulation," Measurement **182**, 109678 (2021).
[28]D. Ellis, "tensorflow/models," https://github.com/tensorflow/models/tree/master/research/audioset/vggish (Last viewed December 2021).
[29]K. L. Hansen, P. Nguyen, G. Micic, B. Lechat, P. Catcheside, and B. Zajamšek, "Amplitude modulated wind farm noise relationship with annoyance: A year-long field study," J. Acoust. Soc. Am. **150**(2), 1198–1208 (2021).
[30]J. Jakobsen, "Infrasound emission from wind turbines," J. Low Freq. Noise Vib. Act. Control **24**(3), 145–155 (2005).
[31]Amplitude Modulation Working Group, "Discussion document: Methods for rating amplitude modulation in wind turbine noise," Institute of Acoustics, St. Albans, UK, 2015.
[32]I. T. Jolliffe, *Principal Component Analysis for Special Types of Data* (Springer, New York, 2002).
[33]L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction," arXiv:1802.03426 (2018).
[34]T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA (August 13–17, 2016), pp. 785–794.
[35]J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," Adv. Neural Inform. Process. Syst. **25**, 2960–2968 (2012).
[36]R Foundation, "The R Project for Statistical Computing," http://www.r-project.org/ (Last viewed December 2021).
[37]P. D. Nguyen, "Long-term quantification of wind farm noise—Audio data" (2021), https://figshare.com/articles/dataset/dataset_for_JASA_express_letter/17451830/2 (Last viewed December 2021).
[38]D. P. Nguyen, "ducphucnguyen/TransferLearningWFN," https://github.com/ducphucnguyen/TransferLearningWFN (Last viewed December 2021).
[39]M. Krzywinski, "Circos," http://circos.ca/ (Last viewed December 2021).
[40]N. L. D. Khoa, M. M. Alamdari, T. Rakotoarivelo, A. Anaissi, and Y. Wang, "Structural health monitoring using machine learning techniques and domain knowledge based features," in *Human and Machine Learning* (Springer, New York, 2018), pp. 409–435.
[41]A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," Renew. Energy **133**, 620–635 (2019).