Presence of 1/*f* noise in the temporal structure of psychoacoustic parameters of natural and urban sounds

Ming Yang
School of Architecture, University of Sheffield, Sheffield, United Kingdom

Bert De Coensel
Department of Information Technology, Ghent University, Ghent, Belgium

Jian Kang[a)]
School of Architecture, University of Sheffield, Sheffield, United Kingdom

1/*f* noise or pink noise, which has been shown to be universal in nature, has also been observed in the temporal envelope of music, speech, and environmental sound. Moreover, the slope of the spectral density of the temporal envelope of music has been shown to correlate well to its pleasing, dull, or chaotic character. In this paper, the temporal structure of a number of instantaneous psychoacoustic parameters of environmental sound is examined in order to investigate whether a 1/*f* temporal structure appears in various types of sound that are generally preferred by people in everyday life. The results show, to some extent, that different categories of environmental sounds have different temporal structure characteristics. Only a number of urban sounds considered and birdsong, generally, exhibit 1/*f* behavior on short to medium duration time scales, i.e., from 0.1 s to 10 s, in instantaneous loudness and sharpness, whereas a more chaotic variation is found in birdsong at longer time scales, i.e., of 10 s–200 s. The other sound categories considered exhibit random or monotonic variations in the different time scales. In general, this study shows that a 1/*f* temporal structure is not necessarily present in environmental sounds that are commonly perceived as pleasant.
© 2015 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4927033]

[GB] Pages: 916–927

I. INTRODUCTION

Pink noise, flicker noise, or "1/*f*" noise is a form of temporal fluctuation (not restricted to noise signals) that has a power density inversely proportional to the frequency (that is, varies as 1/*f*) (Gilden *et al.*, 1995). 1/*f* noise exhibits a predictability that is intermediate between white noise, a quantity with a $1/f^0$ spectral density, and brown noise, a quantity with a $1/f^2$ spectral density. White noise is uncorrelated with its past, showing rapid and random changes; $1/f^2$ noise depends very strongly on its past, showing only slow and predictable changes; 1/*f* noise has an intermediate behavior with some correlation over all time scales, yet not depending too strongly on its past, and exhibits a balance between randomness and correlation on all time scales (Voss and Clarke, 1975; Voss, 1979).

1/*f* noise has been ubiquitously detected in the temporal behavior of systems as diverse as vacuum tubes (Johnson, 1925), river flow (Mandelbrot and Wallis, 1969), traffic flow (Nagatani, 1993), ecological time series (Steele, 1985; Pimm and Redfearn, 1988), DNA-base sequences (Voss, 1992), and nerve membrane (Lundström and McQueen, 1974). The universality of 1/*f* noise is generally believed to be a manifestation of self-organized criticality (SOC) of complex system (Bak *et al.*, 1987).

In the domain of acoustics, 1/*f* behavior has been observed in the power spectral density of the temporal envelope of music and speech (Voss and Clarke, 1978) and of outdoor soundscapes (De Coensel *et al.*, 2003) over different time scales. While the spectral density of the audio signal (i.e., instantaneous sound pressure) of music is usually far from 1/*f*, the spectral density of more slowly varying quantities, i.e., the instantaneous audio power, sound pressure level (varying closely with loudness), or fundamental frequency (pitch) of many musical pieces has been found to vary over time approximately as 1/*f* noise. The spectral density of the instantaneous audio power of speech (as recorded from radio stations) has also been shown to resemble 1/*f* noise, although for pitch (zero crossing rate), a slightly different behavior was found (Voss and Clarke, 1978; De Coensel *et al.*, 2003). Voss and Clarke (1978) explained the 1/*f*-like spectral density of quantities associated with music and speech as the result of a critical balance between predictability and novelty. This was later reinterpreted as music being an imitation of the temporal fluctuation of SOC systems that seems to be common in nature (De Coensel *et al.*, 2003).

For many outdoor soundscapes, consisting of a mixture of sounds originating from multiple sources, De Coensel *et al.* (2003) found the presence of 1/*f* behavior in the *A*-weighted sound level, Zwicker's loudness, and pitch, at least to some extent. By splitting the time interval of interest at around a few seconds, they could distinguish between time structure at the micro-scale and time structure at the macro-scale

a)Electronic mail: j.kang@sheffield.ac.uk

(Botteldooren *et al.*, 2006). At the micro-scale, corresponding to temporal fluctuations in the time interval between 200 ms and 5 s, typically associated to variations within an acoustic event (within sources), a close $1/f$ frequency dependence was often observed in the power spectra of loudness and pitch of various rural and urban soundscapes. However, at the macro-scale, corresponding to temporal fluctuations in the time interval between 5 s and ~10 min, mainly influenced by fluctuations in the succession of acoustic events (between sources), the power spectral density of loudness and pitch of all the rural soundscapes and some of the urban soundscapes that were considered was found to be steeper than $1/f$. This implies slower variations in loudness and pitch than expected in the case of SOC, which indicates a higher predictability (De Coensel *et al.*, 2003).

For music, the slope of the spectral density of the temporal envelope has been shown to correlate well to its pleasing, dull, or chaotic character. Stochastic musical compositions, in which the frequency and duration of each note were determined by $1/f$ noise, have been found to sound pleasing (Voss and Clarke, 1978). Stochastic compositions generated by white noise sources were found to sound too random, chaotic, and unpredictable, whereas those generated by $1/f^2$ noise sounded too correlated, predictable, and hence boring and dull (Voss and Clarke, 1978; De Coensel *et al.*, 2003). Furthermore, reports provide evidence for tuning to $1/f$ dynamics in auditory cortical neurons, which may contribute to the efficient encoding of audio signals for which the temporal envelope shows a $1/f$ behavior (Garcia-Lazaro *et al.*, 2006).

It is well known that certain types of environmental sounds, such as water sounds and birdsong, are well preferred in everyday outdoor sound environments (Yang and Kang, 2005; Nilsson and Berglund, 2006; Axelsson *et al.*, 2010; Kang and Zhang, 2010; Brambilla *et al.*, 2013). Since a common preference for music with a $1/f$ temporal structure has been shown (Voss and Clarke, 1978), this paper aims to explore whether a $1/f$ temporal structure could be observed in a range of sounds that are commonly denoted as being pleasant. This work does not attempt to investigate directly the link between these temporal structure indicators and human perception (such as pleasant, dull, or chaotic), but explores whether a $1/f$ temporal structure is present in distinct categories of environmental sounds that are commonly perceived as being pleasant. Also, this work explores whether the $1/f$ temporal structure indicators have the ability to show differences between different types of sound, and thus contribute to the automatic distinction/identification of various types of environmental sound. In contrast to earlier research (De Coensel *et al.*, 2003) in which mixtures of sounds are considered, in this work, the sound signals of separate sound sources are examined. Next to this, a wider range of psychoacoustic parameters is considered, including an analysis of the temporal structure of the envelope within separate audio frequency bands.

In Sec. II, the methods of sound sample collection, psychoacoustic, $1/f$, and statistical analyses are presented. In Sec. III, the results of an analysis of the temporal structure of single-valued psychoacoustic parameters, including loudness, sharpness, tonality, pitch, and pitch strength, are presented. In Sec. IV, the results of an analysis of the temporal structure of specific loudness are presented. Finally, in Sec. V, the results are discussed.

II. METHODS

A. Sound collection

To study the presence of $1/f$ behavior within the temporal structure of the sound of a particular source, only sound recordings in which a single sound source is predominantly present are used in this study. Sounds that are often heard in everyday life are considered, including sounds from both nature and human activity/facility (Brown *et al.*, 2011). Natural sounds considered include water sounds (stream, river, and sea waves), wind sounds (in deciduous/coniferous trees and in heathland), and birdsongs (in woodland, heathland, grassland, moorland, wetland, farmland, and coastal). Human related sounds considered include sounds of church bells, fountains, street music, street machines (e.g., cleaning machine, rubbish bins loading, and construction work), traffic, human voice, and footsteps in urban areas.

In order to obtain a large sample of single source sound, the recordings used in this study were collected from multiple databases, supplemented by recordings made by the authors. Recordings were made in countryside, natural parks, and urban areas in England from 1994 to 2010. Calibration was based on measured sound pressure levels, or in case calibration data was not available, based on reasonable estimates of the sound level range (Yang, 2013; Yang and Kang, 2013b). In total, 102 single-channel recordings with a sample rate of 44.1 kHz were considered, each having a duration ranging from 30 s to 240 s, with the majority having a duration of 240 s. The recordings can be further classified into four categories: water sounds, wind sounds, birdsong, and urban sounds; the number of recordings in each category can be found in Table I. More details on the set of recordings considered can be found in Yang and Kang (2013b).

B. Psychoacoustic analysis

The temporal variation of a number of psychoacoustic parameters used in previous soundscape analyses (Genuit and Fiebig, 2006; Yang and Kang, 2013a,b), including loudness, sharpness, tonality, pitch, and pitch strength, has been analyzed for each of the 102 selected recordings. Also, the potential $1/f$ structure of the fluctuation of the sound power is examined within each critical band separately. For this purpose, the fluctuation of specific loudness is considered, which is related to the amplitude of oscillations of a particular part of the basilar membrane or the neural responses in the auditory system; similar results could be expected if the root-mean-square value or the amplitude envelope would be considered (Yang, 2013). The calculation of loudness, specific loudness, sharpness, and tonality over time was made using the ArtemiS 11 software package of HEAD Acoustics GmbH (2011a). The calculation of pitch and pitch strength was made using the MIR Toolbox, a MATLAB toolbox dedicated to the extraction of musically related features from

J. Acoust. Soc. Am. **138** (2), August 2015

Yang *et al.* 917

TABLE I. Descriptives of 1/*f* structure indices of loudness, sharpness, tonality, and pitch for the four categories.

		Water				Wind		Bird		Urban	
		Stream		Sea							
Number of recordings		14		20		23		28		17	
Number of recordings with duration of 240 s		10		18		22		23		2	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
N	*N* slope (0.005–10 Hz)	−0.15	0.08	−1.31	0.16	−1.64	0.12	−0.81	0.15	—	—
	N slope (1–10 Hz)	−0.30	0.25	−1.16	0.45	−0.50	0.40	−1.13	0.41	−1.05	0.55
	N slope (0.1–1 Hz)	−0.04	0.25	−2.41	0.73	−2.05	0.37	−1.21	0.62	−1.11	0.80
	N slope (0.005–0.1 Hz)	−0.15	0.42	−0.39	0.33	−1.51	0.45	−0.50	0.45	−0.99	0.43
	N deviation (0.005–10 Hz)	0.36	0.05	0.58	0.11	0.47	0.08	0.45	0.07	—	—
	N deviation (1–10 Hz)	0.08	0.02	0.09	0.01	0.08	0.01	0.11	0.03	0.19	0.07
	N deviation (0.1–1 Hz)	0.22	0.03	0.24	0.03	0.21	0.03	0.24	0.03	0.31	0.09
	N deviation (0.005–0.1 Hz)	0.23	0.07	0.25	0.05	0.24	0.06	0.26	0.05	0.25	0.05
S	*S* slope (0.005–10 Hz)	−0.13	0.12	−1.08	0.14	−1.45	0.22	−0.86	0.16	—	—
	S slope (1–10 Hz)	−0.19	0.09	−0.96	0.30	−0.63	0.25	−1.01	0.41	−0.87	0.51
	S slope (0.1–1 Hz)	−0.04	0.28	−2.14	0.51	−1.61	0.55	−1.37	0.49	−1.07	0.59
	S slope (0.005–0.1 Hz)	−0.22	0.37	−0.15	0.51	−1.43	0.48	−0.51	0.45	−1.00	0.08
	S deviation (0.005–10 Hz)	0.39	0.06	0.56	0.11	0.43	0.06	0.44	0.07	—	—
	S deviation (1–10 Hz)	0.09	0.03	0.08	0.01	0.08	0.02	0.11	0.03	0.19	0.08
	S deviation (0.1–1 Hz)	0.25	0.02	0.24	0.04	0.22	0.02	0.24	0.03	0.28	0.07
	S deviation (0.005–0.1 Hz)	0.27	0.07	0.28	0.06	0.25	0.06	0.25	0.06	0.21	0.01
Ton	Ton slope (0.05–10 Hz)	−0.01	0.06	−0.06	0.08	−0.07	0.07	−0.50	0.14	−0.59	0.33
	Ton deviation (0.05–10 Hz)	0.29	0.04	0.27	0.03	0.26	0.03	0.31	0.04	0.40	0.10
P	PV slope (0.05–10 Hz)	−0.03	0.17	−0.42	0.22	−0.31	0.22	−0.60	0.29	−0.46	0.36
	PA slope (0.05–10 Hz)	−0.09	0.16	−0.42	0.26	−0.27	0.25	−0.59	0.28	−0.65	0.28
	PV deviation (0.05–10 Hz)	0.37	0.04	0.35	0.06	0.33	0.04	0.38	0.05	0.40	0.07
	PA deviation (0.05–10 Hz)	0.34	0.04	0.37	0.05	0.34	0.04	0.35	0.06	0.39	0.08
N	(0.1–1 s)	Random		1/*f* behavior		Random		1/*f* behavior		1/*f* behavior	
	(1–10 s)	Random		Predictable		Predictable		1/*f* behavior		1/*f* behavior	
	(10–200 s)	Random		Random		Predictable		Random		—	
S	(0.1–1 s)	Random		1/*f* behavior		Random		1/*f* behavior		1/*f* behavior	
	(1–10 s)	Random		Predictable		Predictable		Predictable		1/*f* behavior	
	(10–200 s)	Random		Random		Predictable		Random		—	
Ton	(0.1–20 s)	—		—		—		Random		Random	
PV	(0.1–20 s)	Random		Random		Random		Random		Random	
PA	(0.1–20 s)	Random		Random		Random		Random		Random	

audio recordings within the context of music information retrieval (MIR) (Lartillot and Toiviainen, 2007). In the following paragraphs, more detailed information about the calculation of each of the above mentioned parameters is given.

For loudness, the calculation was based on ISO 532B (ISO, 1975; Zwicker *et al.*, 1984) and the computer program specified in the German standard DIN 45631 (Deutsches Institut für Normung e.V., 1991; HEAD Acoustics GmbH, 2011b). The specific loudness pattern for calculating the total loudness was calculated on the basis of 1/3-octave band levels. The 1/3-octave band levels over time were calculated through a fast Fourier transform (FFT) with a window length of 92.88 ms (4096 samples), a Hanning window, and an overlap of 50%. Free field conditions were assumed in the loudness calculation in this study.

Sharpness, a dimension of timbre correlated to the distribution of spectral energy (von Bismarck, 1974), was calculated based on the algorithm of Aures (1985), which takes into account the influence of absolute loudness upon sharpness. As the sharpness calculation algorithm is based on the distribution of specific loudness along the frequency scale, the same procedure as for the loudness calculation was used.

Tonality, another dimension of timbre that indicates whether a sound consists mainly of tonal components or broadband noise, was calculated according to the method of Aures (1985), following Terhardt's procedure for the extraction of tonal components (Terhardt *et al.*, 1982). Again, the same FFT parameters were used in the ArtemiS calculation.

Pitch may be defined as "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" (American Standards Association, 1960; Moore, 1997). Independent of pitch value, pitch strength specifies the pitch sensation along a scale from faint to distinct pitch (Zwicker and Fastl, 1999). Whereas in both the studies of Voss and Clarke (1978) and De Coensel *et al.* (2003), instantaneous pitch was estimated by the rate of zero crossings of the audio signal, in this paper, instantaneous pitch and pitch

strength were calculated according to a simplified model of pitch perception (Meddis and Hewitt, 1991; Moore, 1997; Yang, 2013) implemented in MATLAB (Yang and Kang, 2013a). For the calculation of fluctuation of pitch over time, a window length of 46.4 ms and a hop length of 10 ms were used.

The calculation of specific loudness for Sec. IV was based on German standard DIN 45631/A1 (Deutsches Institut für Normung e.V., 2010), which describes the calculation of time-dependent loudness. This method differs from the total loudness calculation method used above since the DIN standard had not been released at the time when the total loudness calculation was performed. For the analysis of specific loudness, 19 critical bands were considered, covering the frequency range from ∼90 to 15 500 Hz; lower frequency critical bands were not included since a number of the recordings were low-cut filtered.

It should be noted that there might be different algorithms from those used in this study, although currently there are only a limited number of relevant standards. However, since the relative differences are not significant for fluctuations over time (Fastl *et al.*, 2009), it can reasonably be expected that the exact calculation method (including the temporal solutions as discussed following in Sec. III A) will not influence the main conclusions of this study significantly. For the remainder of the paper, the following shorthand is used: N for loudness, S for sharpness, Ton for tonality, PV for pitch value, and PA for pitch strength.

C. Analysis of 1/*f* temporal structure

The temporal structure of instantaneous loudness, sharpness, tonality, pitch, pitch strength, and specific loudness was analyzed using the algorithm outlined in Botteldooren *et al.* (2006) and De Coensel and Botteldooren (2006). First, the spectral density of each of the above parameters is estimated using an FFT over the full duration of each time series. Subsequently, the spectral density is locally averaged over a symmetric frequency interval considering 24 points on a logarithmic scale (De Coensel *et al.*, 2003). Finally, descriptors for the 1/*f* structure are extracted by fitting a straight line through the spectral density plotted on a log –log scale; in particular, the slope and the quadratic deviation from the best-fitted straight line are considered. A slope of -1 and a deviation of 0 would correspond to perfect 1/*f* behavior. Both descriptors can be seen as statistic properties of the instantaneous psychoacoustic parameters, in addition to commonly used indicators, such as mean, standard deviation, and percentile values (Yang and Kang, 2013b).

D. Statistical analysis

A number of statistical methods, including one-way analysis of variance (ANOVA) and one-sample *t*-test, were applied to explore the presence of 1/*f* temporal behavior (slope and deviation indices) in the fluctuations of the various psychoacoustic parameters considered. For this, the SPSS Statistics software package was used (SPSS, 2009).

One-way ANOVA was used to compare the means among the sound categories, to examine if the categories differ from each other significantly in the temporal structure of one or more parameters. The analysis of variance first tests the hypothesis that all group means are equal (*F*-test), and then involves a more detailed examination of the differences if significant differences among the means have been suggested (Hilton and Armstrong, 2006; SPSS, 2009). To compare the mean of the sample of the different sound types in terms of the slope indices to -1, a one-sample *t*-test was used in which the Student's *t*-statistic calculates the ratio of the mean difference (the difference between the sample mean and the test value of -1) to the standard error of the mean.

III. TEMPORAL STRUCTURE OF BASIC PSYCHOACOUSTIC PARAMETERS

A. Frequency range of spectral density

For the majority of the recordings that have a duration of 240 s, the 1/*f* behavior is analyzed over the frequency range (0.005–10 Hz), which corresponds to the time range of 0.1–200 s, according to the duration of the recordings and the temporal resolution of the calculation of the fluctuation of the parameters considered. For the calculation of loudness, sharpness, and tonality as described above, the same window length was used, i.e., 92.9 ms with 50% overlap. For pitch and pitch strength, a window length of 46.4 ms with hop length of 10 ms was used. Thus, the temporal constants and window lengths used in this paper are shorter than the time range considered for the spectral density analyses and, therefore, do not affect significantly the results. Recordings with a duration <240 s are not analyzed for the full frequency range due to the duration limitation.

Examination of the spectral density of variations of loudness, sharpness, tonality, pitch value, and pitch strength with time for the 102 recordings shows that the slope for the different environmental sounds varies considerably. For example, in Fig. 1, which shows several typical shapes of the spectral density of loudness variation with time, the spectrum curve is often not a straight line, but shows breaks at different points. Breaks in the curves occur between -1.0 and $-0.5 \log(\text{Hz})$ in Fig. 1(b), at about $0.0 \log(\text{Hz})$ in Fig. 1(c), and at both ∼-1.0 and $0.0 \log(\text{Hz})$ in Fig. 1(e); in Figs. 1(a) and 1(d), the spectrum curves are, generally, a straight line, although with different slopes.

In order to describe the shapes and slopes of spectral density more precisely, it proved to be useful to further divide the analyzed frequency range into several sub-ranges. For loudness and sharpness, breaks often occur at points of ∼-1.0 and $0.0 \log(\text{Hz})$. The full analyzed frequency range of (0.005–10 Hz) is thus divided into three ranges, i.e., (0.005–0.1 Hz), (0.1–1 Hz), and (1–10 Hz) [corresponding to $(-2.3, -1.0)$, $(-1.0, 0.0)$, and $(0.0, 1.0)$ in logarithmic scale in the figures], which corresponds to the time ranges of (10–200 s), (1–10 s), and (0.1–1 s), respectively. In each of the frequency ranges, the slope of spectral density and its corresponding deviation are calculated. Consequently, eight indices are derived from the spectral density for both loudness and sharpness, that is, slopes and deviations of frequency ranges of (0.005–0.1 Hz), (0.1–1 Hz), (1–10 Hz), and
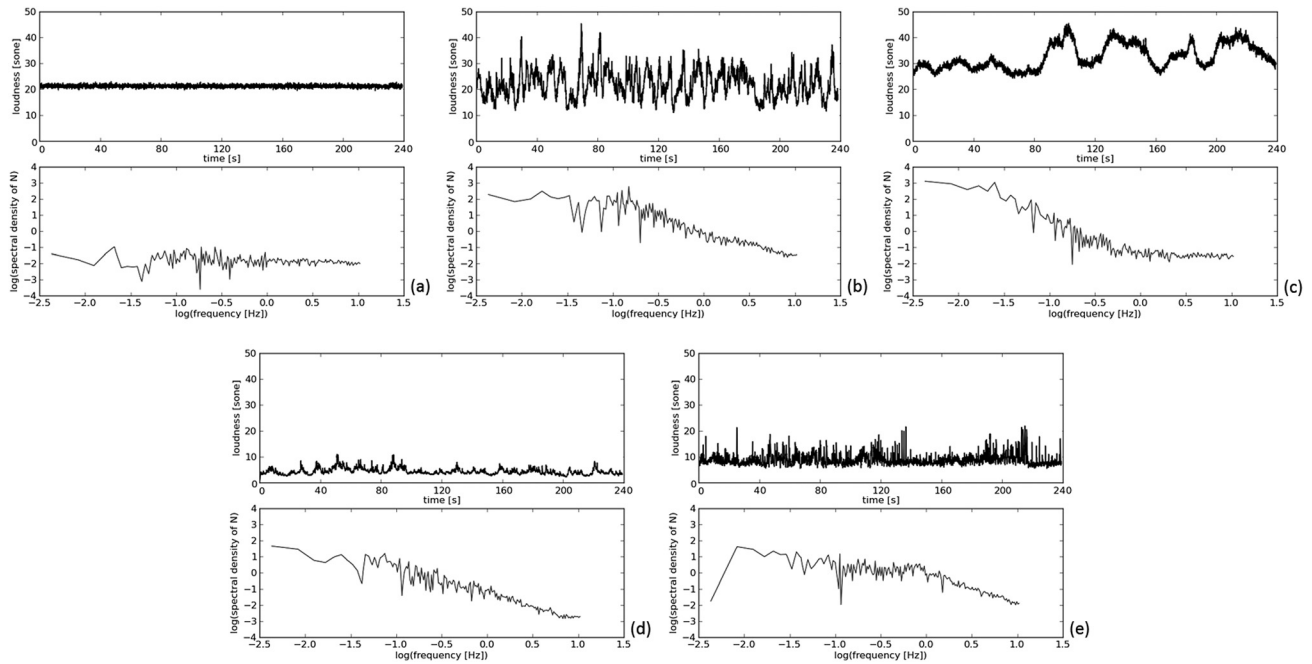
FIG. 1. Time series of loudness, and spectral density of loudness, respectively, corresponding to the sound recording of (a) a river, (b) sea waves, (c),(d) two examples of the sound of wind, and (e) the sound of birdsong.

(0.005–10 Hz) as a whole. For recordings with a duration of 240 s (75 recordings in total), the 4 frequency ranges are analyzed; for recordings with a duration <240 s, only the 2 frequency ranges (0.1–1 Hz) and (1–10 Hz) are analyzed.

In contrast to the spectral density of loudness and sharpness variations, the spectral density of tonality, pitch value, and pitch strength generally turns out to be relatively straight without break points. Thus, single frequency ranges can be used to describe the spectrum curves for these indicators. Considering the duration limits for all the 102 recordings (30 s for shortest ones), the frequency range of (0.05–10 Hz) is used for the analysis of the 1/f structure of tonality, pitch value, and pitch strength, which corresponds to the time range of 0.1–20 s. As a result, only two indices are derived for these parameters, i.e., slope and deviation of spectral density in the frequency range of (0.05–10 Hz).

All the 1/f structure indices for analyses are shown in Table I. From Table I, it can be seen that the mean deviations of both loudness and sharpness of the full range of (0.005–10 Hz) are larger (~1.5–7 times) than those of the

three sub-ranges; that is, it proves the effectiveness of the three sub-ranges in the analysis. Thus, in the following analyses for both loudness and sharpness, only slope and deviation indices of the three sub-ranges are used.

B. Correlations between 1/f temporal structure indices of the psychoacoustic parameters

The Pearson's correlations between the 1/f temporal structure indices of the psychoacoustic parameters as discussed above are analyzed for the 102 recordings; part of the results (slope indices) is shown in Table II, where high correlation coefficients (>0.7) are highlighted in bold numbers. The results show that the 1/f temporal structure indices of sharpness and those of loudness are highly correlated. That is, slopes of sharpness in the three different frequency ranges have high correlations with those of loudness in the same range. This could be expected because both are calculated on the basis of specific loudness; only the weighting over the frequencies differs. Also, there are high correlations between PV slope (0.05–10 Hz) and PA slope (0.05–10 Hz).

TABLE II. Correlations of 1/f structure indices of psychoacoustic parameters.

	N slope (1–10 Hz)	N slope (0.1–1 Hz)	N slope (0.005–0.1 Hz)	S slope (1–10 Hz)	S slope (0.1–1 Hz)	S slope (0.005–0.1 Hz)	Ton slope (0.05–10 Hz)	PV slope (0.05–10 Hz)	PA slope (0.05–10 Hz)
N slope (1–10 Hz)	1.000								
N slope (0.1–1 Hz)	−0.005	1.000							
N slope (0.005–0.1 Hz)	−0.336	0.270	1.000						
S slope (1–10 Hz)	**0.828**	0.162	−0.131	1.000					
S slope (0.1–1 Hz)	0.140	**0.900**	0.204	0.239	1.000				
S slope (0.005–0.1 Hz)	−0.307	0.059	**0.737**	−0.135	0.024	1.000			
Ton slope (0.05–10 Hz)	0.468	−0.128	−0.146	0.387	−0.008	−0.041	1.000		
PV slope (0.05–10 Hz)	0.439	0.264	−0.042	0.461	0.349	−0.095	0.426	1.000	
PA slope (0.05–10 Hz)	0.555	0.108	−0.101	0.543	0.184	−0.164	0.592	**0.742**	1.000

C. Comparison of 1/f structure indices among sound categories

The significant mean differences among the sound categories considered in the 1/f temporal structure indices are examined using an ANOVA. Levene's test of homogeneity of variances of the four different categories shows that for N slope (0.005–0.1 Hz), N deviation (0.005–0.1 Hz), S slope (0.005–0.1 Hz), S deviation (0.005–0.1 Hz), and PA slope (0.05–10 Hz), the assumption of homogeneity of variance has been met (p values are greater than the α level of 0.05), that is, the variances of different categories are equal; for the other indices, the variances are unequal. The results of ANOVA show that for the majority of indices, except for N deviation (0.005–0.1 Hz), S slope (0.1–1 Hz), S deviation (0.005–0.1 Hz), and PA deviation (0.05–10 Hz), the p values associated with the F ratio (the ratio of between-groups variance to within-groups variance) are less than the α level of 0.05, suggesting there might be some statistically significant differences between the means of at least two categories. However, since the assumption of homogeneity of variances has not been met for a number of indices, *post hoc* tests are further checked for verifying and examining which of the specific categories differ.

According to the homogeneity of variances of the categories of each index, different methods are used to produce *post hoc* tests: Tukey's honestly significant difference test and Dunnett's T3 method are used, respectively, for indices with equal and unequal variances between categories, considering the unequal sample sizes in groups (Stevens, 1999). The mean and standard deviation of the indices for the four categories are shown in Table I. The results of the *post hoc* tests, shown in Table III, present the pairwise differences among means of the categories.

In the following sections (Secs. III D–III G), the 1/f structure dynamic characteristics of the four sound categories and their differences are discussed in terms of the four different parameters.

D. 1/f structure of loudness fluctuations

To analyze the 1/f characteristics of the four sound categories in terms of loudness, the results for the 102 recordings are plotted in the two-dimensional coordinate system with axes representing the slope and deviation indices of spectral density of loudness in the different ranges, as shown in Fig. 2. In the figures, which display the ranges of (0.005–0.1 Hz) and (0.005–10 Hz), that is Figs. 2(a), 2(d), and 2(f), only two urban sounds are shown, including a music and a traffic sound, since only these have a duration of 240 s and thus are analyzed in the ranges in the urban category. From the plots and, in particular, from Figs. 2(a) and 2(c), it can be seen that the water sounds can be classified into two groups. Detailed data analysis showed that the recordings of stream and river sounds in the water sound category are in one of the groups, while the recordings of sea wave sound are in the other. Also, while the recordings in water, wind, and bird categories are, generally, apart from each other in Figs. 2(e) and 2(f), the recordings in the urban category are more dispersive and mixed with many of the recordings in the other three categories.

As shown in Table III, the slopes over the ranges of (1–10 Hz) [N slope (1–10 Hz)] and (0.1–1 Hz) [N slope (0.1–1 Hz)] show significant differences between the categories of wind and bird sound, and between wind and urban sound categories. From Table I, it can be seen that for N slope (1–10 Hz), the mean values are -0.3 to -0.5 for stream and river sounds and wind sounds, and -1.1 for sea waves, bird, and urban sounds. For N slope (0.1–1 Hz), the mean values are 0.0 for stream and river sounds, -2.1 to -2.4 for sea waves and wind sounds, and -1.1 to -1.2 for bird and urban sounds. The slope over the range of

TABLE III. Multiple comparisons of 1/f structure indices of loudness, sharpness, tonality, and pitch for the four categories, where * indicate significantly different group means at the level of 0.05.

(I) Category (J) Category	Mean Difference (I - J) Dependent Variable	Water			Wind			Bird			Urban		
		Wind	Bird	Urban	Water	Bird	Urban	Water	Wind	Urban	Water	Wind	Bird
N	N slope (1–10 Hz)	−0.31	0.33	0.25	0.31	0.64*	0.56*	−0.33	−0.64*	−0.08	−0.25	−0.56*	0.08
	N slope (0.1–1 Hz)	0.62	−0.22	−0.33	−0.62	−0.84*	−0.94*	0.22	0.84*	−0.10	0.33	0.94*	0.10
	N slope (0.005–0.1 Hz)	1.20*	0.20	—	−1.20*	−1.00*	—	−0.20	1.00*	—	—	—	—
	N deviation (1–10 Hz)	0.01*	−0.03*	−0.11*	−0.01*	−0.04*	−0.12*	0.03*	0.04*	−0.08*	0.11*	0.12*	0.08*
	N deviation (0.1–1 Hz)	0.02	−0.01	−0.07*	−0.02	−0.03*	−0.10*	0.01	0.03*	−0.06	0.07*	0.10*	0.06
	N deviation (0.005–0.1 Hz)	0.00	−0.01	—	0.00	−0.01	—	0.01	0.01	—	—	—	—
S	S slope (1–10 Hz)	−0.01	0.36*	0.23	0.01	0.37*	0.24	−0.36*	−0.37*	−0.13	−0.23	−0.24	0.13
	S slope (0.1–1 Hz)	0.34	0.10	−0.21	−0.34	−0.24	−0.54*	−0.10	0.24	−0.30	0.21	0.54*	0.30
	S slope (0.005–0.1 Hz)	1.26*	0.33	—	−1.26*	−0.93*	—	−0.33	0.93*	—	—	—	—
	S deviation (1–10 Hz)	0.00	−0.02*	−0.10*	0.00	−0.03*	−0.11*	0.02*	0.03*	−0.08*	0.10*	0.11*	0.08*
	S deviation (0.1–1 Hz)	0.03*	0.01	−0.03	−0.03*	0.02	−0.06*	−0.01	0.02	−0.04	0.03	0.06*	0.04
	S deviation (0.005–0.1 Hz)	0.02	0.03	—	−0.02	0.01	—	−0.03	−0.01	—	—	—	—
Ton	Ton slope (0.05–10 Hz)	0.03	0.46*	0.55*	−0.03	0.43*	0.52*	−0.46*	−0.43*	0.09	−0.55*	−0.52*	−0.09
	Ton deviation (0.05–10 Hz)	0.02	−0.03	−0.12*	−0.02	−0.04*	−0.13*	0.03	0.04*	−0.09*	0.12*	0.13*	0.09*
P	PV slope (0.05–10 Hz)	0.05	0.34*	0.20	−0.05	0.29*	0.15	−0.34*	−0.29*	−0.14	−0.20	−0.15	0.14
	PA slope (0.05–10 Hz)	−0.01	0.31*	0.37*	0.01	0.32*	0.38*	−0.31*	−0.32*	0.06	−0.37*	−0.38*	−0.06
	PV deviation (0.05–10 Hz)	0.03	−0.02	−0.04	−0.03	−0.05*	−0.07*	0.02	0.05*	−0.02	0.04	0.07*	0.02
	PA deviation (0.05–10 Hz)	0.01	0.00	−0.04	−0.01	−0.01	−0.05	0.00	0.01	−0.04	0.04	0.05	0.04

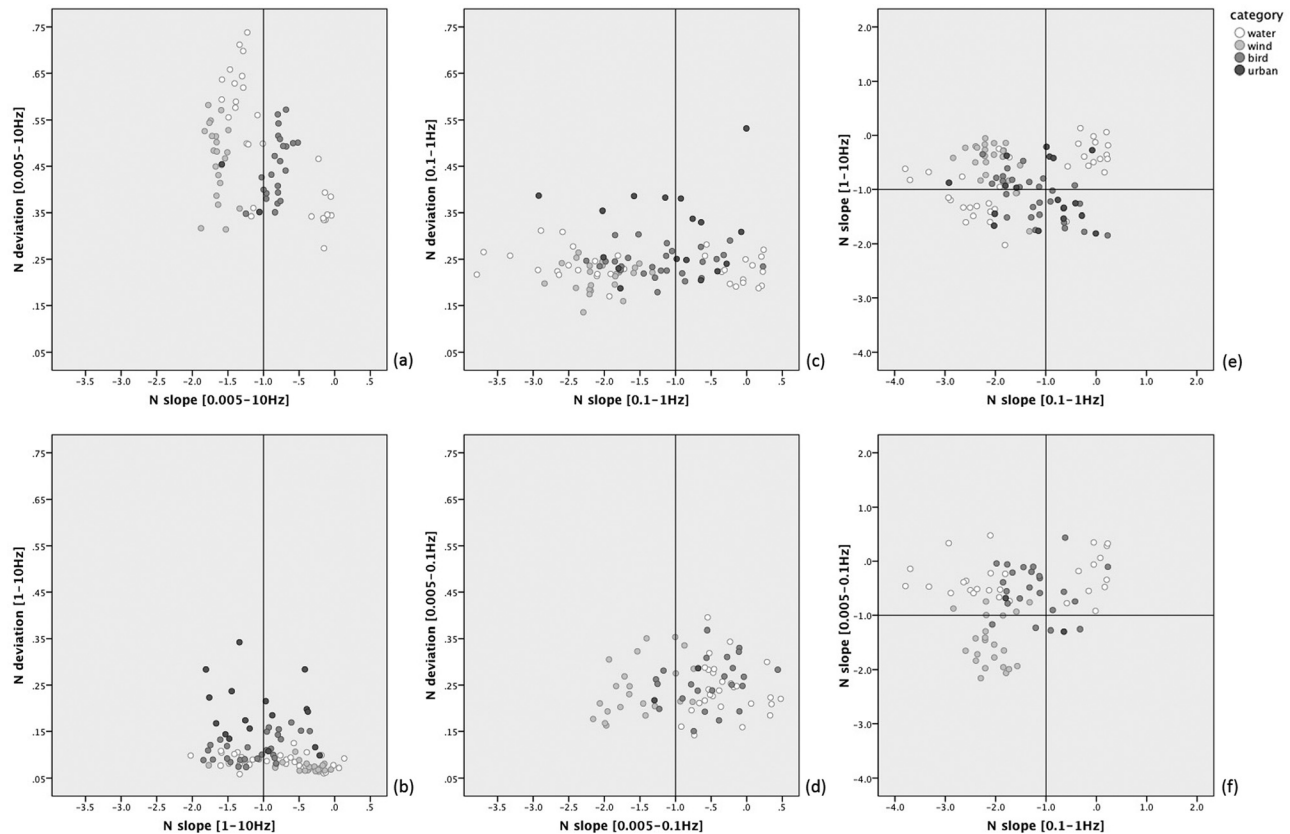J. Acoust. Soc. Am. **138** (2), August 2015

Yang *et al.* 921

FIG. 2. Characteristics of the four categories of sound in terms of the $1/f$ structure of instantaneous loudness.

(0.005–0.1 Hz) [N slope (0.005–0.1 Hz)] shows significant differences between wind and water, and between wind and birdsongs. The mean values for N slope (0.005–0.1 Hz) are -1.5 for wind category and -0.3 to -0.5 for water and bird categories. In terms of deviations from the slopes of spectra of loudness, in the ranges of (1–10 Hz) and (0.1–1 Hz), the mean deviation values of the urban category are generally higher than those of the other three categories, while deviation of the range of (0.005–0.1 Hz) does not show any significant difference between the different categories.

From the results above, it can be summarized that the means of slopes of spectral density of loudness over the different ranges mainly distinguish wind sounds from bird and urban categories of sounds. Stream and river sounds in the water sound category exhibit relatively random variations (slopes of 0.0 to -0.3) in loudness compared to $1/f$ behavior in all the three frequency ranges, which correspond to 0.1–200 s. Sea wave sounds exhibit a $1/f$ behavior (slope of -1.2) in loudness fluctuations in the short time range of 0.1–1 s, more predictable variations (slope of -2.4) in the medium time range of 1–10 s, and more random variations (slope of -0.4) in the long time range of 10–200 s. Wind sounds exhibit relatively random variations (slope of -0.5) compared to $1/f$ behavior on short time range, and more predictable variations (slopes of -1.5 to -2.1) on medium and long time ranges. Birdsong sounds generally exhibit a $1/f$ behavior (slopes of -1.1 to -1.2) in loudness on short and medium time ranges, and more random variations (slope of -0.5) on long time range. Urban sounds generally exhibit $1/f$ behavior (slopes of -1.1) on short and medium time ranges.

Among urban sounds, in particular, for which the standard deviations of slopes are high (0.6–0.8), speech sounds exhibit a $1/f$ behavior in loudness fluctuations in the short time range and more random variations in the medium time range. Sound from traffic at a distance exhibits random variations in both the short and medium time ranges, whereas close traffic exhibits a $1/f$ behavior in the short time range and more predictable variations in the medium time range. Some street machines exhibit $1/f$ behavior in the short time range, while some exhibit $1/f$ behavior in the medium time range. Fountains exhibit relatively random variations (slopes of -0.2 to -0.4) in the short time range similar to stream and river sounds and $1/f$ behavior or predictable variations (slopes of 1.0 to -1.8) in the medium time range. Church bells exhibit predictable variations (slopes of -1.8) in the short time range and $1/f$ behavior or random variations (slopes of 0.0 and -1.1) in the medium time range. Footsteps exhibit random variations (slope of -0.4) in the short time range, similarly, and $1/f$ behavior (slope of -0.9) in the medium time range. Street music in this study does not show $1/f$ behavior in any of the time ranges, rather they exhibit predictable variations (slopes of -1.5) in the short time range and random variations (slopes of -0.3 and -0.6) in the medium time range. Among birdsongs in the medium time range, for which the standard deviation of slope is also high (0.6), the subcategories do not show any distinct temporal structure.

Overall, for water sounds and birdsongs, which are generally perceived as pleasant in an outdoor context, only sea wave sounds and birdsongs in a limited time range show $1/f$
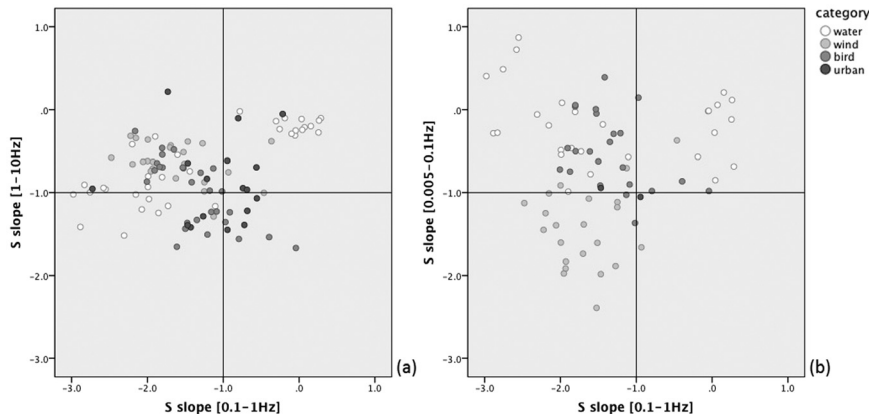
922   J. Acoust. Soc. Am. **138** (2), August 2015

Yang *et al.*

FIG. 3. Characteristics of the four categories of sound in terms of the 1/*f* structure of instantaneous sharpness.

behavior. Mechanical and traffic sounds, which are generally perceived as unpleasant, also show 1/*f* behavior in some time ranges. That is, a 1/*f* temporal structure is not necessarily present within environmental sounds that are commonly perceived as pleasant.

Between water and wind sound categories, stream and river sounds and wind sounds have similar random variations in the range of (1–10 Hz); sea wave sounds and wind sounds have similar predictable variations in the range of (0.1–1 Hz). However, the whole water sound category differs from the wind sound category in the range of (0.005–0.1 Hz). While water and wind sounds have been shown to have similar psychoacoustic characteristics, making them not easy to be distinguished (Yang and Kang, 2013b), they are quite different in terms of the spectral density of temporal fluctuations.

E. 1/*f* structure of sharpness fluctuations

Since the 1/*f* temporal structure indices of sharpness and those of loudness are highly correlated as discussed above in Sec. III B, the results of sharpness and those of loudness show a similar 1/*f* structure for the different sound categories. The 1/*f* characteristics of the four categories of sounds in terms of sharpness can be seen in Fig. 3, where the recordings in different sound categories are located at a similar location as compared to those based on loudness, shown in Figs. 2(e) and 2(f), although they are somewhat more mixed based on sharpness. That is, the 1/*f* behavior of sharpness generally resembles that of loudness.

F. 1/*f* structure of tonality fluctuations

From Table III, it can be seen that the slope of spectral density of tonality [Ton slope (0.05–10 Hz)] shows significant differences between the water and wind sound categories and the bird and urban sound categories. As shown in Table I, the mean slope values are 0.0 to −0.1 for water and wind categories, and −0.5 to −0.6 for bird and urban categories. The mean deviation value [Ton dev. (0.05–10 Hz)] of the urban category is higher than those of the other three categories. The 1/*f* characteristics of the four sound categories in terms of the slope and deviation indices of tonality are also presented in Fig. 4(a). Overall, in the time interval of 0.1 s–20 s, bird and urban sounds exhibit relatively random variations in tonality compared to 1/*f* behavior, while for water and wind sounds, which generally do not show any tonality (i.e., with tonality values around zero) (Yang and Kang, 2013b), the mean slopes of spectral density are equal to about zero.

G. 1/*f* structure of pitch fluctuations

Analysis of the slope and deviation of the spectral density of pitch value and pitch strength fluctuations over the range of (0.05–10 Hz) generally reveals significant differences between the water and wind sound categories and the bird and urban sound categories, similar to the results of tonality, as shown in Table III. In Table I, for both pitch value and pitch strength, the mean slopes are ~−0.3 for water and wind categories, and −0.5 to −0.7 for bird and urban
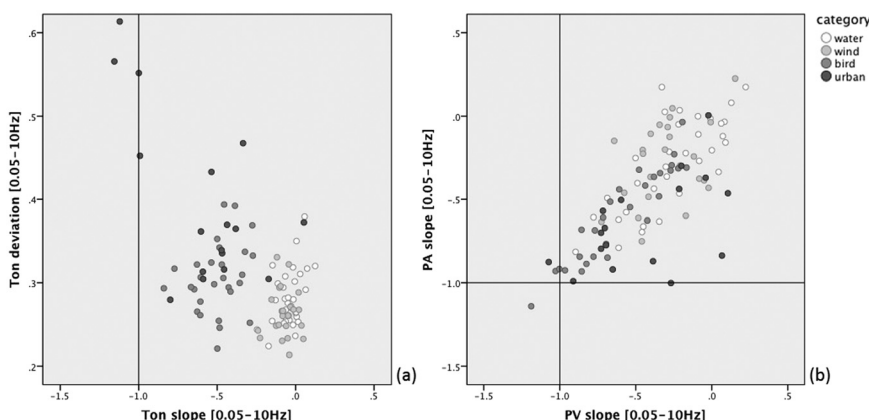

FIG. 4. Characteristics of the four categories of sound in terms of the 1/*f* structure of (a) tonality and (b) pitch.

J. Acoust. Soc. Am. **138** (2), August 2015

Yang *et al.* 923

categories. There are, generally, no significant differences between the categories regarding the deviations of the spectral density. Figure 4(b) shows the characteristics of the four categories in two dimensions. In the time interval of 0.1 s–20 s, bird and urban sounds exhibit relatively random variations in both pitch value and pitch strength, while for water and wind sounds, the mean slopes of the spectral density are even closer to zero than those of bird and urban sounds, and exhibit random variations as well.

IV. TEMPORAL STRUCTURE OF SPECIFIC LOUDNESS

A. Frequency range of spectral density

Over the frequency range of (0.005–1 Hz), excluding the range of (1–10 Hz) that was considered in the loudness analysis in Sec. III C, considering the time internal of available data of specific loudness (0.40 s), the shapes of the spectral density of specific loudness for the 102 recordings show breaks often occurring at points of $\sim -1.0 \log(\text{Hz})$. Therefore, similar to the case of total loudness, the $1/f$ structure of specific loudness is analyzed over two frequency ranges separately, i.e., (0.1–1 Hz) and (0.005–0.1 Hz), which correspond to the medium and long time internal ranges of (1–10 s) and (10–200 s). In each frequency range, the slope of the spectral density and its corresponding deviation are calculated for each of the 19 critical bands. That is, 76 indices in total are derived for further analysis. The recordings with durations < 240 s are analyzed only in the frequency range of (0.1–1 Hz), and recordings with duration of 240 s are analyzed for both ranges.

B. Correlations between the $1/f$ temporal structure indices

The correlations between the 76 $1/f$ temporal structure indices for specific loudness are first examined. It is found that within a certain frequency range, correlations between the slope indices are generally high, among which the correlation coefficients are > 0.8 between the slopes of adjacent three to seven critical bands in the range of (0.1–1 Hz), and between the slopes of adjacent two to four critical bands in the range of (0.005–0.1 Hz). The correlations between the deviation indices within each frequency region are not high, generally < 0.6 for both the frequency ranges. Relatively high correlations (~ 0.5–0.8) exist between adjacent critical bands. Correlations between slopes across the two different frequency ranges, between the deviation indices across the two frequency ranges, and between slope and deviation indices within or across the frequency ranges, are not significant, with correlation coefficients of, generally, < 0.4. In addition, the correlations between the $1/f$ temporal structure indices of specific loudness and those of total loudness (as in Sec. III C) are examined; part of the results is shown in Table IV. The results show that both the slope and deviation indices of specific loudness of the 19 critical bands have certain correlations with those of total loudness only within each frequency range with correlation coefficients of ~ 0.5–0.8.

As discussed above, within each frequency range, the variations of specific loudness in adjacent critical bands

	N slope (1–10 Hz)	N slope (0.1–1 Hz)	N slope (0.005–0.1 Hz)
N3 slope (0.1–1 Hz)	0.318**	0.672**	−0.131
N8 slope (0.1–1 Hz)	0.305**	0.693**	−0.035
N13 slope (0.1–1 Hz)	0.258**	**0.830****	−0.052
N18 slope (0.1–1 Hz)	0.347**	0.641**	−0.316**
N3 slope (0.005–0.1 Hz)	−0.256*	0.274*	**0.706****
N8 slope (0.005–0.1 Hz)	−0.225	0.268*	**0.771****
N13 slope (0.005–0.1 Hz)	−0.290*	0.141	**0.781****
N18 slope (0.005–0.1 Hz)	−0.163	0.166	**0.722****

show similar slopes and also somewhat similar deviations of spectral density; the number of adjacent bands that have high correlations depends on the frequency range. Also, the variations of specific loudness have certain correlations with those of total loudness within each frequency range. Thus, to analyze the characteristics of different categories of sound in terms of variation of specific loudness, only part of the slope indices is presented in the following sections (as well as in Table IV), which includes the slopes of 3rd, 8th, 13th, and 18th bands in both the frequency ranges of (0.1–1 Hz) and (0.005–0.1 Hz).

C. $1/f$ structure of specific loudness fluctuations

In order to examine whether variations of specific loudness in each of the frequency bands considered exhibit $1/f$ behavior, and whether this $1/f$ behavior is more pronounced than in the case of total loudness, a one-sample t-test is used here. Table V shows the mean and standard deviation of the slope indices of each sound type, where asterisks indicate the population means that are not significantly different from the hypothesized value of -1. From the results, it can be concluded that in the frequency range of (0.1–1 Hz), the temporal variation of specific loudness in some critical bands of wind sounds and birdsongs, and of specific loudness in all critical bands of urban sounds, shows $1/f$ behavior, as well as the temporal variation of total loudness of birdsongs and urban sounds in this frequency range.

For wind sounds, while temporal variation of total loudness in the range of (0.1–1 Hz) does not show $1/f$ behavior, temporal variations of specific loudness in certain critical bands do. For birdsong sounds, while temporal variation of total loudness does show $1/f$ behavior in this frequency range, this is only the case for specific loudness within a few critical bands. However, the results do not allow the conclusion that the variations of specific loudness are more likely to exhibit $1/f$ behavior than those of total loudness.

D. Comparison of $1/f$ structure indices for specific loudness among sound categories

With the eight slope indices as discussed above in Sec. IV B, the sound recordings are plotted in the two-dimensional coordinate system shown in Fig. 5. From the

TABLE V. Statistics of one-sample test (test value = −1) of 1/*f* structure indices of specific loudness and loudness, where * indicate the population mean is not significantly different from the hypothesized value at the 0.05 level.

	Stream		Sea		Wind		Bird		Urban	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
N slope (0.1–1 Hz)	−0.044	0.252	−2.407	0.730	−2.052	0.370	**−1.210***	0.616	**−1.108***	0.799
N slope (0.005–0.1 Hz)	−0.151	0.424	−0.386	0.335	−1.506	0.451	−0.505	0.453	—	—
*N*3 slope (0.1–1 Hz)	−0.020	0.226	−1.637	0.623	−0.692	0.438	−0.478	0.331	**−0.962***	0.789
*N*8 slope (0.1–1 Hz)	0.067	0.249	−1.724	0.718	**−0.865***	0.361	−0.619	0.554	**−0.803***	0.655
*N*13 slope (0.1–1 Hz)	−0.097	0.194	−1.885	0.702	**−1.069***	0.490	**−1.030***	0.477	**−1.054***	0.652
*N*18 slope (0.1–1 Hz)	−0.002	0.198	−1.811	0.675	−0.511	0.444	−1.256	0.448	**−0.793***	0.778
*N*3 slope (0.005–0.1 Hz)	−0.136	0.541	−0.404	0.391	−1.356	0.539	**−0.841***	0.465	—	—
*N*8 slope (0.005–0.1 Hz)	−0.010	0.188	−0.302	0.425	−1.410	0.459	−0.567	0.585	—	—
*N*13 slope (0.005–0.1 Hz)	−0.080	0.519	−0.172	0.415	−1.473	0.411	−0.524	0.498	—	—
*N*18 slope (0.005–0.1 Hz)	0.086	0.504	−0.137	0.429	−1.264	0.405	−0.560	0.521	—	—

results in both Fig. 5 and Table V, it can be seen that in the frequency range of (0.005–0.1 Hz), for all of the three natural sound categories, i.e., water, wind, and bird sounds, the characteristics in 1/*f* structure (in terms of slope values) do not differ much with different frequency bands, and also do not differ much from those of total loudness discussed in Sec. III C. In the range of (0.1–1 Hz), however, while for water (including both stream and river recordings and sea wave recordings) and urban sounds, the behavior of specific loudness in different critical bands and of total loudness does not differ much, for wind sounds and birdsongs, the behavior varies with different bands and differs from those of total loudness. For example, for wind sounds, the mean slopes are −0.5 to −0.7 for the 3rd and 18th bands and close to −1.0 for the 8th and 13th bands in specific loudness, and −2.1 in total loudness. For urban sounds, the slope values of specific loudness of the different critical bands and of total loudness are similar, all close to −1, though the ranges of slope values are relatively wide compared to wind sounds and birdsongs (the standard deviations are 0.7–0.8 for urban sounds, and 0.3–0.6 for wind sounds and birdsongs as shown in Table V).

V. DISCUSSION AND CONCLUSIONS

In this paper, the statistical characteristics of different categories of environmental sounds have been explored in terms of the 1/*f* structure of the spectral density of a range of psychoacoustic parameters, including loudness, sharpness, tonality, pitch value, and pitch strength, as well as specific loudness, i.e., loudness within each critical band.

The environmental sounds considered in this paper show different slopes of the spectral density of loudness over different temporal ranges. The results for sharpness fluctuations show similar tendencies to those of loudness; however, the differences among the categories are not that clear. For both loudness and sharpness, stream and river sounds and sea wave sounds in the water sound category show clear differences in terms of the spectral density slope. Stream and river sounds exhibit relatively random variations (slopes of 0.0 to −0.3) compared to 1/*f* behavior in the full time interval of 0.1 s–200 s; sea wave sounds exhibit, generally,

1/*f* behavior (slopes of −1.0 to −1.2) in the short time range of 0.1–1 s, predictable variations (slopes of −2.1 to −2.4) in the medium time range of 1–10 s, and random variations (slopes of −0.2 to −0.4) in the long time range of 10–200 s. Wind sounds exhibit relatively random variations (slopes of −0.5 to −0.6) compared to 1/*f* behavior in short time range, and more predictable variations (slopes of −1.4 to −2.1) in medium and long time ranges. Birdsongs exhibit generally 1/*f* behavior (slopes of −1.0 to −1.2) in short and medium time ranges (1/*f* behavior on only short time range in sharpness), and more random variations (slopes of −0.5) in long time range. Urban sounds generally exhibit 1/*f* behavior (slopes of −0.9 to −1.1) in short and medium time ranges. While the results for water, wind, and bird sound categories are generally clearly distinct, the results for sounds within the urban category tend to have a much greater spread and overlap with the results of the other three categories.

The slopes of the spectral density of tonality, pitch value, and pitch strength all show significant differences between the water and wind sounds and the bird and urban sounds over the time interval range of 0.1 s–20 s, roughly corresponding to the short and medium time ranges of loudness and sharpness. Bird and urban sounds exhibit relatively random variations in tonality, pitch value, and pitch strength. Water and wind sounds generally do not show any tonality; the mean slopes of the spectral density were found to be close to zero.

When results of specific loudness within critical bands are considered, similar characteristics of 1/*f* structure are found as compared to those found for total loudness for all the sound categories, in general. However, there are some differences, e.g., in the temporal range of 1–10 s, the spectral density characteristics of wind sounds and birdsong sounds vary with critical band and differ from those of total loudness. For wind sounds, temporal variations of specific loudness in certain critical bands do show 1/*f* behavior, though that of total loudness does not.

Generally, these results demonstrated the certain ability of the 1/*f* temporal structure indicators in making the distinction between the environmental sounds; for example, the slope in the range of (0.005–0.1 Hz) shows clear difference
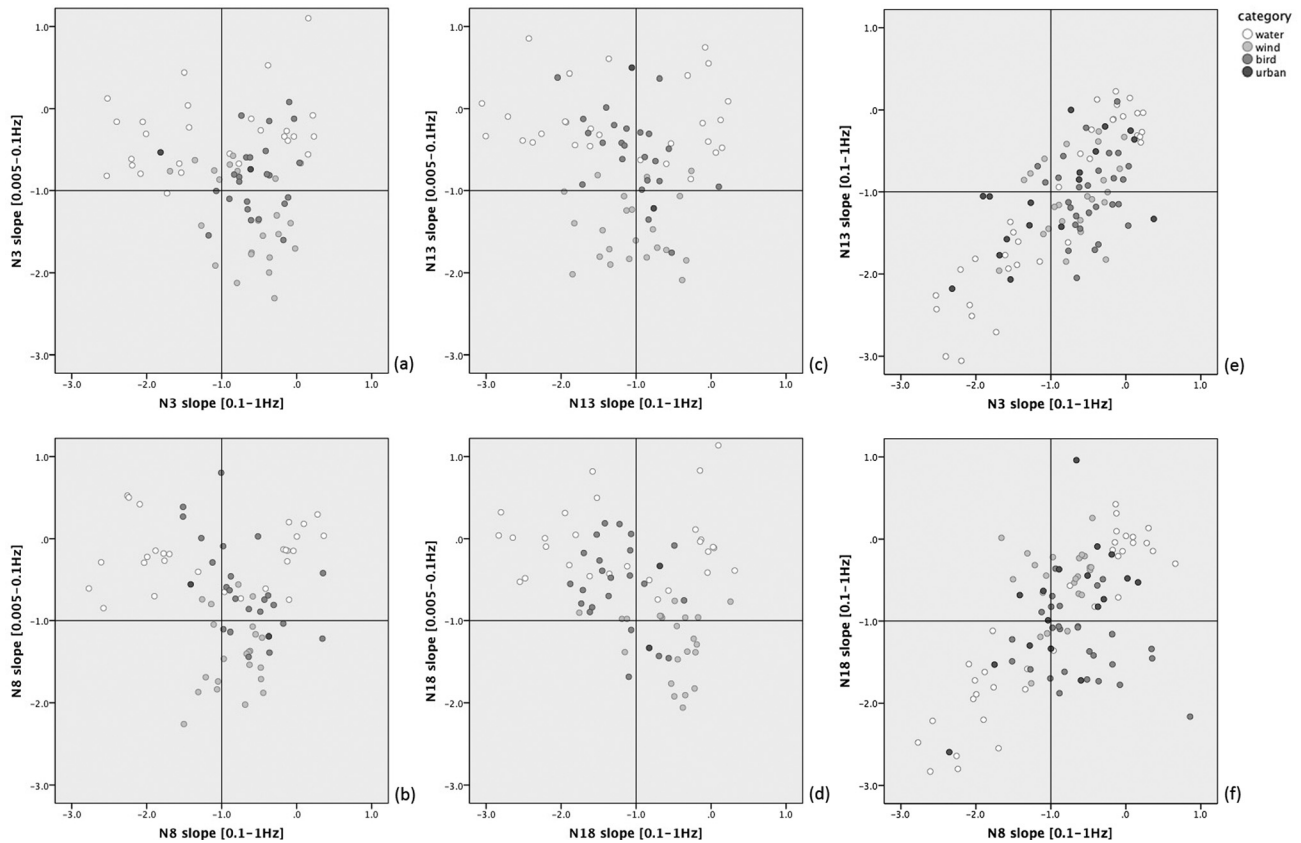
FIG. 5. Characteristics of the four types of sound in terms of the $1/f$ structure slope indices of specific loudness in the ranges of [0.1–1 Hz] and [0.005–0.1 Hz].

between water and wind sounds, which are difficult to be distinguished by psychoacoustic indicators (Yang and Kang, 2013b). However, the $1/f$ temporal structure indicators do not have a direct link with preference of environmental sounds as that in music. The results show that a $1/f$ temporal structure is not necessarily present in environmental sounds that are commonly perceived as pleasant. Among the different categories of sound, only sea wave sounds, birdsongs, and a number of urban sounds (such as speech, traffic, and street machine sounds), generally, exhibit $1/f$ behavior in loudness and sharpness on short and/or medium duration time scales.

These results also confirm the theoretical expectations of $1/f$ noise of loudness and pitch fluctuations in many activities in soundscape (De Coensel *et al.*, 2003), although with some divergences. Based on local wind velocity dynamics (an approximate $1/f$ dependence from seconds to minutes obtained from theoretical considerations and experimental data) and the relation between wind velocity and wind induced noise, a dependence approximating $1/f^2$ was expected for wind induced vegetation sound level (De Coensel *et al.*, 2003). In this study, wind sounds (in vegetation) exhibit slopes of ~-2.0 to -1.5 in loudness in the time interval of 1 s–200 s, which accords with the theoretical considerations. For birdsong, $1/f$ dynamics of loudness and pitch fluctuation were expected (De Coensel *et al.*, 2003) since the necessary ingredients were present to create SOC. In this paper, it is found that birdsongs have slopes of ~-1 in loudness in the time interval of 0.1 s–10 s (but ~-0.5 in the time interval of 10 s–200 s), and slopes of ~-0.6 in pitch in the

time interval of 0.1 s–20 s. The results confirm that birdsong loudness fluctuations exhibit, generally, $1/f$ behavior in short to medium time intervals, but relatively more random variations of pitch.

ACKNOWLEDGMENTS

American Standards Association (**1960**). *Acoustical Terminology SI, 1-1960* (American Standards Association, New York).

Aures, W. (**1985**). "Berechnungsverfahren fur den sensorischen wohlklang beliebiger schallsignale" ("A model for calculating the sensory euphony of various sounds"), Acustica **59**, 130–141.

Axelsson, O., Nilsson, M. E., and Berglund, B. (**2010**). "A principal components model of soundscape perception," J. Acoust. Soc. Am. **128**, 2836–2846.

Bak, P., Tang, C., and Wiesenfeld, K. (**1987**). "Self-organized criticality: An explanation of the $1/f$ noise," Phys. Rev. Lett. **59**, 381–384.

Botteldooren, D., De Coensel, B., and De Muer, T. (**2006**). "The temporal structure of urban soundscapes," J. Sound Vib. **292**, 105–123.

Brambilla, G., Gallo, V., Asdrubali, F., and D'Alessandro, F. (**2013**). "The perceived quality of soundscape in three urban parks in Rome," J. Acoust. Soc. Am. **134**, 832–839.

Brown, A. L., Kang, J., and Gjestland, T. (**2011**). "Towards standardization in soundscape preference assessment," Appl. Acoust. **72**, 387–392.

De Coensel, B., and Botteldooren, D. (**2006**). "The quiet rural soundscape and how to characterize it," Acta. Acust. Acust. **92**, 887–897.

De Coensel, B., Botteldooren, D., and De Muer, T. (**2003**). "1/f noise in rural and urban soundscapes," Acta. Acust. Acust. **89**, 287–295.

Deutsches Institut für Normung e.V. (**1991**). DIN 45631. "Procedure for calculating loudness level and loudness" (Deutsches Institut für Normung e.V., Berlin, Germany).

Deutsches Institut für Normung e.V. (**2010**). DIN 45631/A1. "Calculation of loudness level and loudness from the sound spectrum—Zwicker Method—Amendment 1: Calculation of the loudness of time-variant sound; with CD-ROM" (Deutsches Institut für Normung e.V., Berlin, Germany).

Fastl, H., Völk, F., and Straubinger, M. (**2009**). "Standards for calculating loudness of stationary or time-varying sounds," in *Proceedings of the 38th International Congress and Exposition on Noise Control Engineering*, *INTER-NOISE* 2009, Ottawa, Canada, pp. 1508–1513.

Garcia-Lazaro, J. A., Ahmed, B., and Schnupp, J. W. H. (**2006**). "Tuning to natural stimulus dynamics in primary auditory cortex," Curr. Biol. **16**, 264–271.

Genuit, K., and Fiebig, A. (**2006**). "Psychoacoustics and its benefit for the soundscape approach," Acta. Acust. Acust. **92**, 952–958.

Gilden, D., Thornton, T., and Mallon, M. (**1995**). "1/f noise in human cognition," Science **267**, 1837–1839.

HEAD Acoustics GmbH (**2011a**). "ArtemiS," available at http://www.head-acoustics.de/eng/nvh_artemis.htm (Last viewed July 20, 2013).

HEAD Acoustics GmbH (**2011b**). "Head application note—Psychoacoustic analyses I," available at http://www.headacoustics.de/downloads/eng/application_notes/PsychoacousticAnalysesI_06_11e.pdf (Last viewed July 20, 2013).

Hilton, A., and Armstrong, R. (**2006**). "*Post hoc* ANOVA tests," available at http://eprints.aston.ac.uk/9317/1/Statnote_6.pdf (Last viewed July 20, 2013).

ISO (**1975**). ISO 532:1975B. "Acoustics—Method for calculating loudness level" (International Organization for Standardization, Geneva, Switzerland).

Johnson, J. B. (**1925**). "The Schottky effect in low frequency circuits," Phys. Rev. **26**, 71–85.

Kang, J., and Zhang, M. (**2010**). "Semantic differential analysis of the soundscape in urban open public spaces," Build. Environ. **45**, 150–157.

Lartillot, O., and Toiviainen, P. (**2007**). "A MATLAB toolbox for musical feature extraction from audio," in *Proceedings of the 10th International Conference on Digital Audio Effects*, DAFx-07, Bordeaux, France, pp. 237–244.

Lundström, I., and McQueen, D. (**1974**). "A proposed 1/f noise mechanism in nerve cell membranes," J. Theor. Biol. **45**, 405–409.

Mandelbrot, B. B., and Wallis, J. R. (**1969**). "Some long-run properties of geophysical records," Water Resour. Res. **5**, 321–340, doi:10.1029/WR005i002p00321.

Meddis, R., and Hewitt, M. J. (**1991**). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," J. Acoust. Soc. Am. **89**, 2866–2882.

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing* (Academic, London), pp. 177–212.

Nagatani, T. (**1993**). "Power-law distribution and 1/f noise of waiting time near traffic-jam threshold," J. Phys. Soc. Jpn. **62**, 2533–2536.

Nilsson, M. E., and Berglund, B. (**2006**). "Soundscape quality in suburban green areas and city parks," Acta. Acust. Acust. **92**, 903–911.

Pimm, S. L., and Redfearn, A. (**1988**). "The variability of population densities," Nature **334**, 613–614.

SPSS Inc. (**2009**). Pasw Statistics Base 18, available at http://support.spss.com/ProductsExt/Statistics/Documentation/18/clientindex.html (Last viewed November 19, 2010).

Steele, J. H. (**1985**). "A comparison of terrestrial and marine ecological systems," Nature **313**, 355–358.

Stevens, J. (**1999**). "*Post hoc* tests in ANOVA," available at http://pages.uoregon.edu/stevensj/posthoc.pdf (Last viewed July 20, 2013).

Terhardt, E., Stoll, G., and Seewann, M. (**1982**). "Algorithm for extraction of pitch and pitch salience from complex tonal signals," J. Acoust. Soc. Am. **71**, 671–678.

von Bismarck, G. (**1974**). "Sharpness as an attribute of the timbre of steady sounds," Acustica **30**, 159–172.

Voss, R. F. (**1979**). "1/f (flicker) noise: A brief review," in *Proceedings of the 33rd Annual Symposium on Frequency Control*, Atlantic City, NJ, pp. 40–46.

Voss, R. F. (**1992**). "Evolution of long-range fractal correlations and 1/f noise in DNA base sequences," Phys. Rev. Lett. **68**, 3805–3808.

Voss, R. F., and Clarke, J. (**1975**). " '1/f noise' in music and speech," Nature **258**, 317–318.

Voss, R. F., and Clarke, J. (**1978**). "1/f noise in music: Music from 1/f noise," J. Acoust. Soc. Am. **63**, 258–263.

Yang, M. (**2013**). "Natural and urban sounds in soundscapes," Ph.D. thesis, School of Architecture, The University of Sheffield, Sheffield, UK.

Yang, W., and Kang, J. (**2005**). "Soundscape and sound preferences in urban squares: A case study in Sheffield," J. Urban Des. **10**, 61–80.

Yang, M., and Kang, J. (**2013a**). "Applicability and application of music features in soundscape," in *Proceedings of the 40th Italian Annual Conference on Acoustics (AIA), 39th German Annual Conference on Acoustics (DAGA), combined with European Acoustics Association (EAA) Euroregio (AIA-DAGA)*, Merano, Italy.

Yang, M., and Kang, J. (**2013b**). "Psychoacoustical evaluation of natural and urban sounds in soundscapes," J. Acoust. Soc. Am. **134**, 840–851.

Zwicker, E., and Fastl, H. (**1999**). *Psychoacoustics—Facts and Models* (Springer, Berlin), pp. 203–264.

Zwicker, E., Fastl, H., and Dallmayr, C. (**1984**). "Basic-program for calculating the loudness of sounds from their 1/3-oct band spectra according to ISO 532-B," Acustica **55**, 63–67.

J. Acoust. Soc. Am. **138** (2), August 2015

Yang *et al.*   927