Journal of Sound and Vibration (2002) **250**(1), 9–21
doi:10.1006/jsvi.2001.3884, available online at http://www.idealibrary.com on IDEAL®

DIAGNOSTIC SYSTEM BASED ON THE HUMAN AUDITORY–BRAIN MODEL FOR MEASURING ENVIRONMENTAL NOISE—AN APPLICATION TO RAILWAY NOISE

H. Sakai, T. Hotehama and Y. Ando

Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, *Japan.*
E-mail: sakaih@kobe-u.ac.jp

AND

N. Prodi and R. Pompoli

Department of Engineering, University of Ferrara, Via Saragat 1, 44100 *Ferrara, Italy*

(*Received* 5 *September* 2001)

Measurements of railway noise were conducted by use of a diagnostic system of regional environmental noise. The system is based on the model of the human auditory–brain system. The model consists of the interplay of autocorrelators and an interaural crosscorrelator acting on the pressure signals arriving at the ear entrances, and takes into account the specialization of left and right human cerebral hemispheres. Different kinds of railway noise were measured through binaural microphones of a dummy head. To characterize the railway noise, physical factors, extracted from the autocorrelation functions (ACF) and interaural crosscorrelation function (IACF) of binaural signals, were used. The factors extracted from ACF were (1) energy represented at the origin of the delay, $\Phi(0)$, (2) effective duration of the envelope of the normalized ACF, τ_e, (3) the delay time of the first peak, τ_1, and (4) its amplitude, ϕ_1. The factors extracted from IACF were (5) IACC, (6) interaural delay time at which the IACC is defined, τ_{IACC}, and (7) width of the IACF at the τ_{IACC}, W_{IACC}. The factor $\Phi(0)$ can be represented as a geometrical mean of energies at both ears as listening level, *LL*.

1. INTRODUCTION

Environmental noise is usually evaluated statistically as a sound pressure level (*SPL*) and its frequency characteristic [1]. *SPL* is measured with a sound-level meter. A recently proposed diagnostic system for evaluating environmental noises [2] is based on the model of the human auditory–brain system [3]. As shown in Figure 1, it (1) measures the environmental noise and analyzes its physical factors, (2) identifies the noise source using the extracted physical factors, and (3) evaluates it subjectively. The key feature of the system is that the physical factors to be evaluated are extracted from the autocorrelation functions (ACF) and interaural crosscorrelation function (IACF) for signals arriving at a person's ears in accordance with the auditory–brain model. The receiver is actually a dummy head with binaural microphones as ears so as to reflect human psychology. In the field of concert hall acoustics, such binaural measurements are widely accepted for subjective evaluations based on the auditory–brain model [4]. The physical factors are analyzed at intervals with a short

0022-460X/02/060009 + 13 $35.00/0

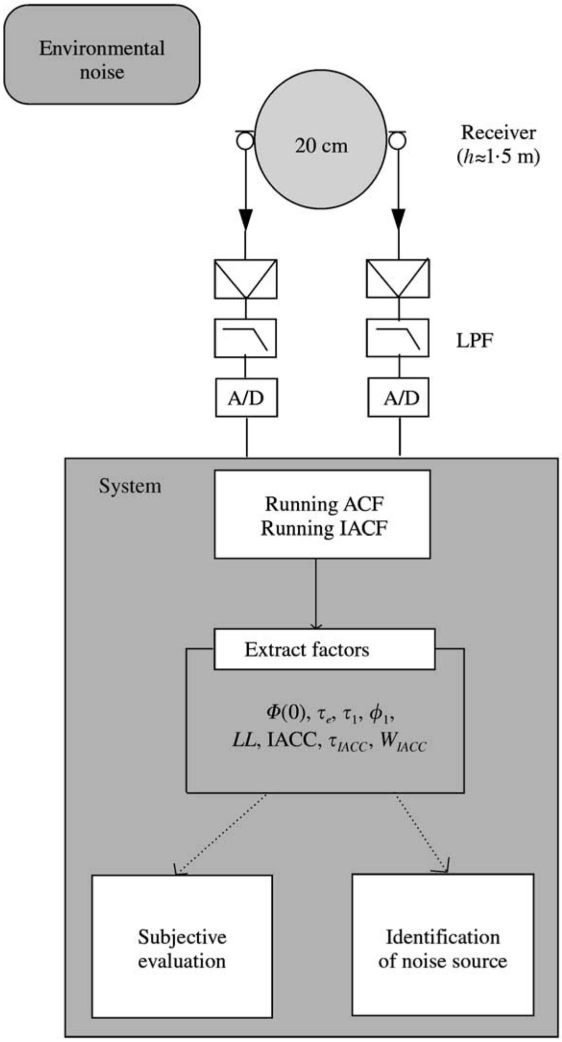

Figure 1. Diagnostic system proposed for measuring, identifying, and subjectively evaluating environmental noise. Physical factors are obtained from binaural signals received at a dummy head. LPF is a low-pass filter.

duration corresponding to the "psychological present [5]" for a noise field. This system has previously been used to characterize aircraft noise [6, 7]. The noise of an aircraft in flight was well characterized. The physical factors are classified as temporal or spatial according to the auditory–brain model, which takes into account the specialization of the left and right cerebral hemispheres as the left and right hemispheres are activated by temporal factors extracted from the ACF and spatial factors extracted from the IACF respectively [4].

In the conventional system, environmental noise is measured monaurally using a sound-level meter, and the physical factors are extracted as a statistics of the sound–pressure levels, like L_x ($x = 5, 50, 95\% \ldots$) or L_{eq}. Such an evaluation system uses spectrum information as the frequency domain. On the other hand, binaural measurements, which are closer to actual human listening conditions, better reflect the many effects of the subjective attributes. This system is also more effective in dealing with environmental noise in the time domain, because the physical factors are extracted from the ACF and IACF of the binaural signals.

The proposed diagnostic system should be able to predict the psychological attributes by using these temporal and spatial factors. For example, a sound may be perceived as noisy in a given situation even though its SPL is quite low. A good example is the beep of a mobile phone in an otherwise quiet train. Thus, loudness is related not only to SPL, but also to τ_e, which is one of the ACF factors (see section 3.2) [4, 8]. Moreover, the phenomenon that the fundamental pitch of a complex tone can be perceived by a person is well known as the "phenomenon of missing fundamental". In this phenomenon, the pitch of harmonic components without a fundamental frequency is perceived as being the same as the pitch of a pure tone of the fundamental frequency. While this fundamental pitch cannot be predicted by frequency analysis of the signal, it can be predicted by temporal analysis even when the complex tone consists of random phase components [9]. With binaural measurement, spatial information including subjective diffuseness and directional information can be obtained from the IACF factors. Consequently, the use of ACF and IACF analysis to subjectively evaluate noise is quite reasonable.

The ACF and IACF factors can also be used to identify a noise source as "timbre" by using multi-dimensional analysis. Recently, the use of four primitive sensations, including duration sensation, was proposed [10]. Duration sensation comparing pure tone and noise varies with τ_e. In addition, a short-time moving (running) ACF and IACF, as described in section 3.1, can be used to evaluate time-variant noise. In addition to the effects for subjective attributes, floor impact noise [11] and transmission loss between rooms [12] can be well described by ACF and IACF factors.

The purpose of the present study was to measure railway noise and characterize it by using the proposed diagnostic system. We investigated the distinctions between different kinds of noise generated by trains and compared the physical factors measured at different distances from the source. We used only acoustical information to evaluate the noise.

2. MEASUREMENTS

2.1. MEASUREMENT LOCATIONS AND MEASURED RAILWAY NOISE

We conducted our noise measurements at two different places–along railway lines in Ferrara (Italy) and along railway lines in Kobe (Japan). Preliminary measurements were conducted in Ferrara to confirm that ACF and IACF factors can represent railway noise (Measurement I). Final measurements were conducted in Kobe using two receivers at different distances from the railway line (Measurement II). The target sound sources were running trains for both cases. The railway noises were recorded manually.

Measurement I was conducted at a fixed single position. The distance between the closest rail and the receiver was about 10 m. There were no large reflective surfaces near the receiver. The noise sources were categorized as: passenger train ("train"), freight train ("freight"), whistling ("whistle"), or steam hissing ("steam"). The total number of trains measured was 31 (train 19; freight 5; whistle 5; and steam 2).

Measurement II was conducted using two different receiver positions. The distances between the closest railway and the two receivers were about 5 (Point 1) and 25 m (Point 2). There were two tracks, and around the receivers there were reflective surfaces (buildings). The noise sources (trains) were categorized as: special express (fastest, $n = 2$), super express (faster, $n = 1$), express ($n = 4$), local ($n = 4$), or freight ($n = 3$) (where "n" represents the number of trains measured). The temperature and humidity were 30°C and 61%, respectively, on average the day we conducted our measurements.

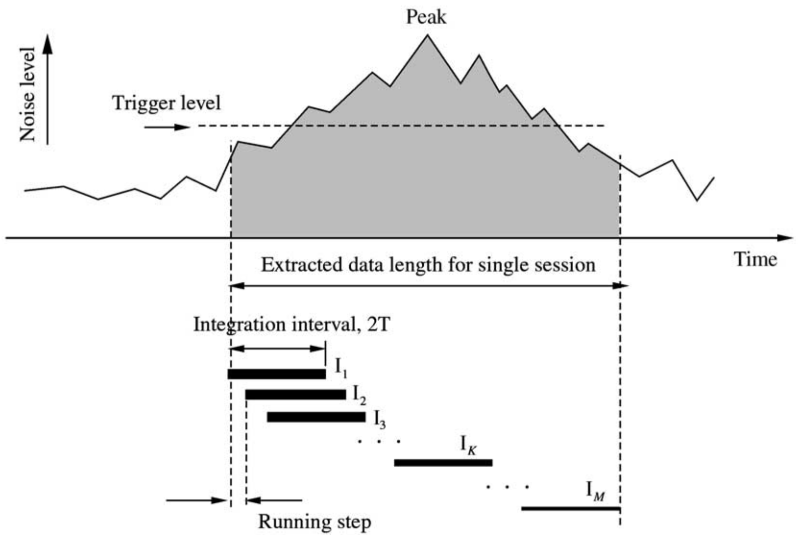

Figure 2. Concept of "running" ACF and IACF in both measurements. Physical factors were obtained from each integration interval, 2T, for each running step.

2.2. PROCEDURES

The measurement system used for both measurements was controlled by a laptop computer (366-MHz CPU; 143-MB RAM) running measurement software developed for this purpose. The noise to be measured was recorded on the computer's hard disk at a sampling frequency of 44·1 kHz. The physical factors were analyzed by the measurement software in real time. Binaural condenser microphones were attached to opposite sides of a sphere (a dummy head) made of styrene foam and having a diameter of 200 mm. The thickness of the foam was 20 mm. The same dummy head was used in both measurements. The microphones were 1·5 m above the ground and aligned parallel to the railway tracks. The concept of a "running" ACF and IACF is illustrated in Figure 2.

The extracted time length for a single session in Measurement I was 5 s. It was set to be 10 s in Measurement II to obtain the longer activity of trains. In the calculation of the running ACF and IACF, the integration interval 2T was 0·5 s, and the running step was 0·1 s. The running ACF was obtained after it passed through an A-weighting network.

3. CALCULATION OF PHYSICAL FACTORS

3.1. PHYSICAL FACTORS IN NOISE FIELDS

The physical factors in the noise fields, described in the next section, were extracted from the ACF and IACF of binaural noise signals. Because environmental noise is not constant, these functions were calculated at certain intervals (integration intervals, 0·5 s), as shown in Figure 2. The start of each integration interval was delayed for a short time (the running step). We used a running step of 0·1 s based on experience. The integration interval can be determined within the "psychological present" (the duration of time that a person feels now). The mid point of the duration was at the center of the maximum $\Phi(0)$, which is one of the ACF factors.

3.2. FACTORS EXTRACTED FROM ACF AND IACF

Orthogonal factors were extracted from the running ACF [8], energy, $\Phi(0)$, the effective duration of the envelope of the normalized ACF, τ_e, the delay time of the first peak, τ_1, and its amplitude, ϕ_1. The $\Phi(0)$ is the sound energy arriving at the ears. This factor is the ACF at the origin of the time delay for each ear, and corresponds to the equivalent SPL. The τ_e is defined by the tenth-percentile delay representing a kind of repetitive feature or reverberation within the source signal itself. The values of τ_e were obtained from the initial regression of the ACF peaks in logarithmic scale, excluding the origin of the ACF. The τ_1 and ϕ_1 are the delay time and amplitude of the first dominant peak of the normalized ACF. The factors ϕ_2, $\phi_3, \ldots$ and τ_2, $\tau_3, \ldots$ are closely related to ϕ_1 and τ_1 respectively. These factors were analyzed for a single ear.

To determine the spatial characteristics of the sound signal, we conducted binaural measurements. From the IACF, we extracted the IACC defined as the maximum value of normalized interaural crosscorrelation function within its delay time ± 1 ms, the interaural delay time at which the IACC is defined, τ_{IACC}, and the width of the IACF at the τ_{IACC}, W_{IACC}. One of the IACF factors, the listening level (LL), can be obtained from the geometrical mean of the monaural $\Phi(0)$-s. The IACC corresponds to subjective diffuseness. The τ_{IACC} is the interaural time delay at the maximum peak; it determines the IACC and corresponds to the horizontal sound localization and the balance of the sound field. In particular, directional information about the noise source can be obtained from this factor. The W_{IACC} is the time interval at the IACF within 10% of the maximum value. It is related to the apparent source width as well as to the IACC [8]. Although $\Phi(0)$ is an ACF factor, the LL is implied as a binaural factor because it is the denominator of the normalized IACF. Thus, LL is the geometrical mean of the sound energies arriving at both ears, $\Phi(0)$:

$$LL = 10 \log[\Phi_{ll}(0)\Phi_{rr}(0)]^{1/2}, \tag{1}$$

where LL is represented on a decibel scale.

4. RESULTS

4.1. MEASUREMENT I

In Measurement I, we compared the physical factors for different types of railway noise. Typical physical factors obtained from the running ACF and IACF for a passenger train are graphed in Figure 3. The monaural ACF factors are shown only for a single ear (left). Not only did the $\Phi(0)$ and LL change, but so did the other ACF and IACF factors. The $\Phi(0)$ and LL are related to the SPL. The τ_{IACC} well describes the directivity of the source at every moment. A running train can be regarded as a linear source, and rapid changes in τ_{IACC} are due to the noise created when the wheels strike the joints of the rails. Changes of the factors for a freight train were almost the same as for a passenger train. The only difference was in τ_{IACC}, which did not change rapidly in the case of a freight train.

Figure 4(a)–(e) shows typical ACF factors and IACC for whistle noise. In $\Phi(0)$, a rapid increase in energy can be seen. Changes in the other factors (τ_e, τ_1, ϕ_1, and IACC) are also quite evident. The τ_e and IACC increased due to the tonal components of the whistling sound. In contrast, τ_e became smaller in the case of steam noise as shown in Figure 4(f). The ϕ_1, which is related to the perception of pitch strength, increased dramatically to almost 1·0.

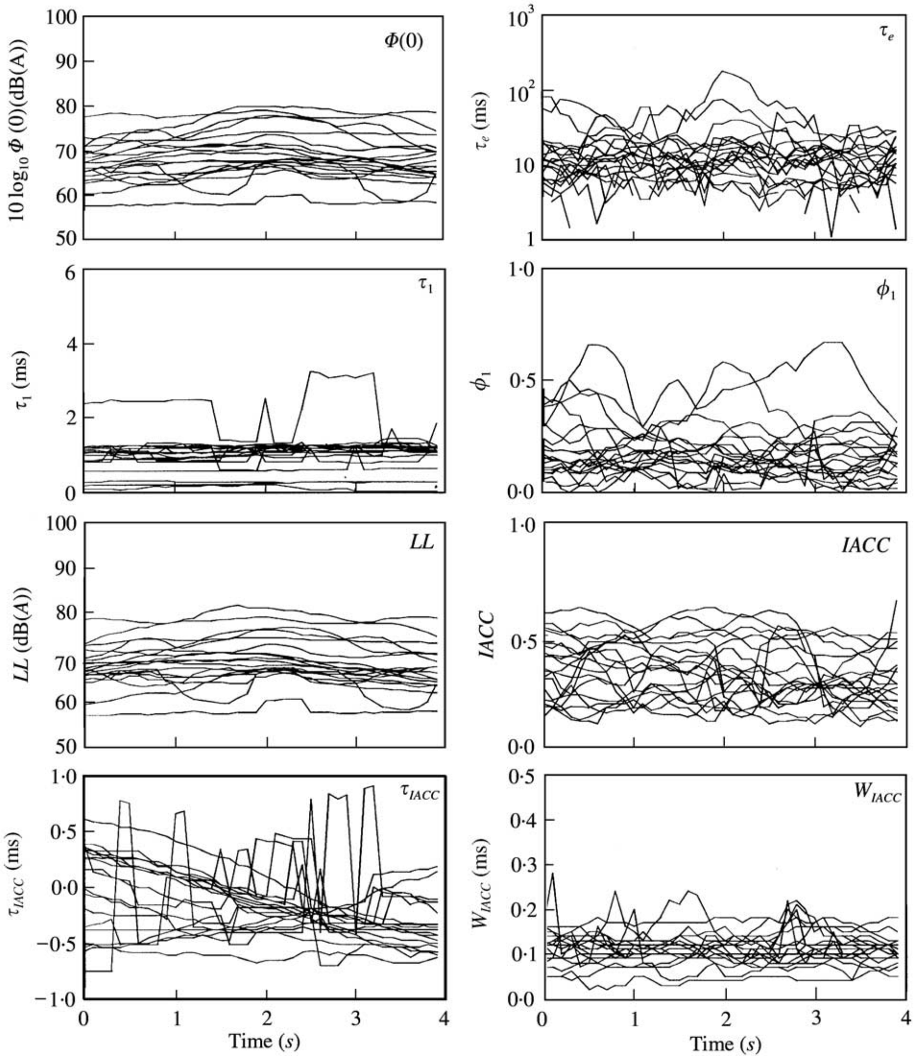

Figure 3. Typical results of all physical factors from running ACF and IACF for passenger train in Measurement I.

4.2. MEASUREMENT II

In Measurement II, we compared the physical factors for the same train measured at two different locations. Typical physical factors obtained from the running ACF and IACF for an express train are graphed in Figure 5. As in Measurement I, not only did *LL* change, but so did the other factors.

Figure 6 shows examples of the measured running τ_e for three different trains. The τ_e for the fastest train (special express) remained small during the measurement, always less than 10 ms. For the local and express trains, it increased during the measurement, especially at Point 2 (further position). The τ_e peaks appeared near the *LL* ones. As shown in Figure 6(b), the intermittent peaks of the running τ_e were generally periodical, with a period of between 2·5 and 3·3 s. This activity can be perceived by a person listening. The waveforms of ACF at

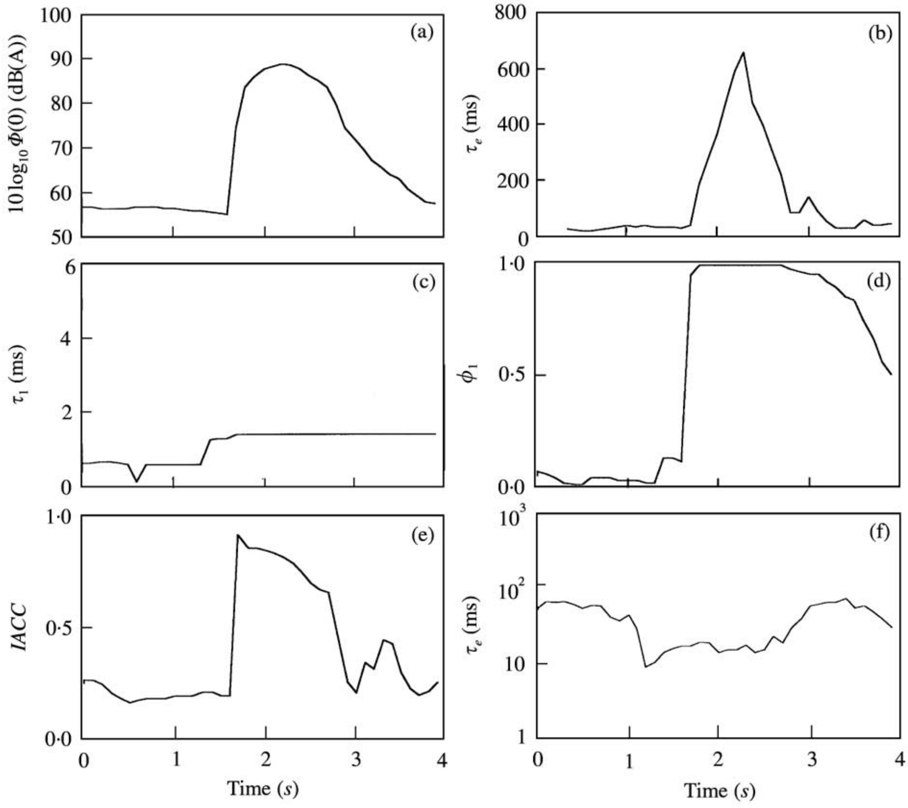

Figure 4. Typical example of whistle noise and steam noise: ACF factors and IACC. (a) running $\Phi(0)$; (b) τ_e; (c) τ_1; (d) ϕ_1 for whistle noise; and (f) running τ_e for steam noise. Note that the vertical axis of (f) is indicated as a logarithmic scale for better explanation.

different times during the measurement and at different positions are shown in Figure 7 in logarithmic scale; they correspond to the example in Figure 6(b). The repetitive feature of the ACF was most prominent at Point 2 for the local train.

The τ_1 and ϕ_1 are related to pitch sensation and its strength respectively. The values of ϕ_1 were less than 0·3 in most cases, meaning that pitch in accordance with τ_1 could not be perceived. Therefore, a person listening perceives railway noise as simply noise without any specific tonal component. In our measurements, the ϕ_1 varied only for the freight train as shown in Figure 8. At 3·8 s, ϕ_1 peaked near 0·4, meaning that the pitch corresponding to τ_1 (around 1 kHz) could be perceived by a person listening.

The IACC temporarily decreased at the *LL* peak at both positions. As shown in Figure 5, the IACC was the same at both positions for the same measurement period.

The values of τ_{IACC} at Points 1 and 2 do not represent the rapid changes in source direction, as there were no obstacles between the railway and receiver after 5 s in Figure 5. This was due to the large reflective surfaces of the buildings. The constant τ_{IACC} values in initial parts were due to the strong reflections from the lateral walls of the buildings on both sides of the receiver. The rapid changes in τ_{IACC} might be due to the noise created when the wheels struck the joints of the rails. This activity was relatively weak at Point 2. Figure 9 shows a typical activity for W_{IACC} for three different trains. Generally, the W_{IACC} at Point 2 was larger than that at Point 1. This is because the effect of the absorption of air was larger at higher frequencies at Point 2.

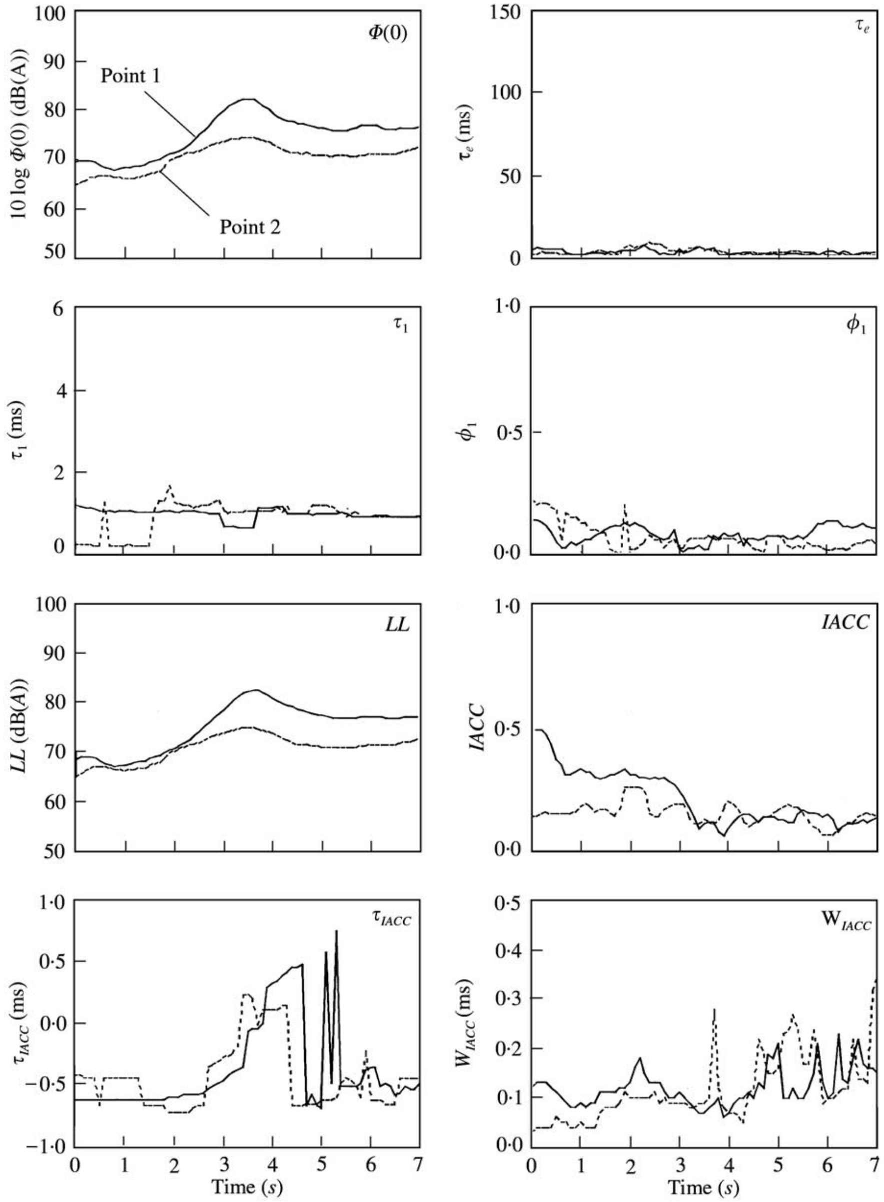

Figure 5. Typical results of all physical factors from running ACF and IACF for express train in Measurement II.

5. DISCUSSION

As described in section 4.1, not only did $\Phi(0)$ and LL vary with different kinds of railway noise, but so did the other ACF and IACF factors. Figure 10 shows the distribution of values obtained at the maximum $\Phi(0)$ for physical factors for the four types of noise sources. Significant differences were obtained between the types for some factors, as shown in Table 1. The τ_e varied greatly due to the tonal or noise component, which is included in each noise source. Thus, τ_e well represents the characteristics of each noise source, as do $\Phi(0)$ and

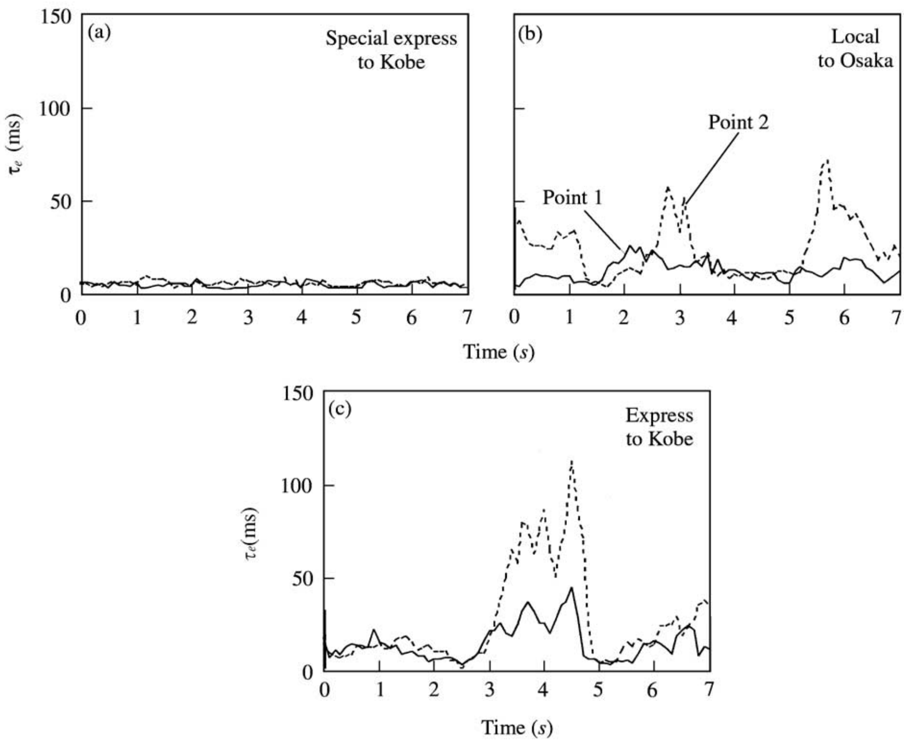

Figure 6. Examples of measured running τ_e for three different trains. Straight and dotted lines represent results at Points 1 and 2 respectively: (a) result of running τ_e for special express to Kobe; (b) result of running τ_e for local to Osaka; (c) result of running τ_e for express to Kobe.

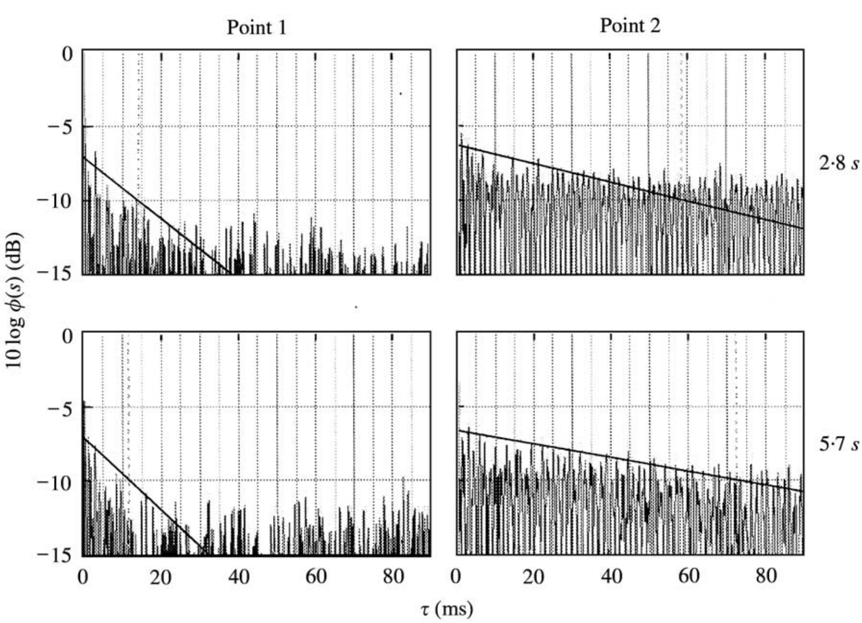

Figure 7. Waveforms of ACF for different times (2·8 and 5·7 s) at different positions in logarithmic scale, corresponding to results shown in Figure 9(b).

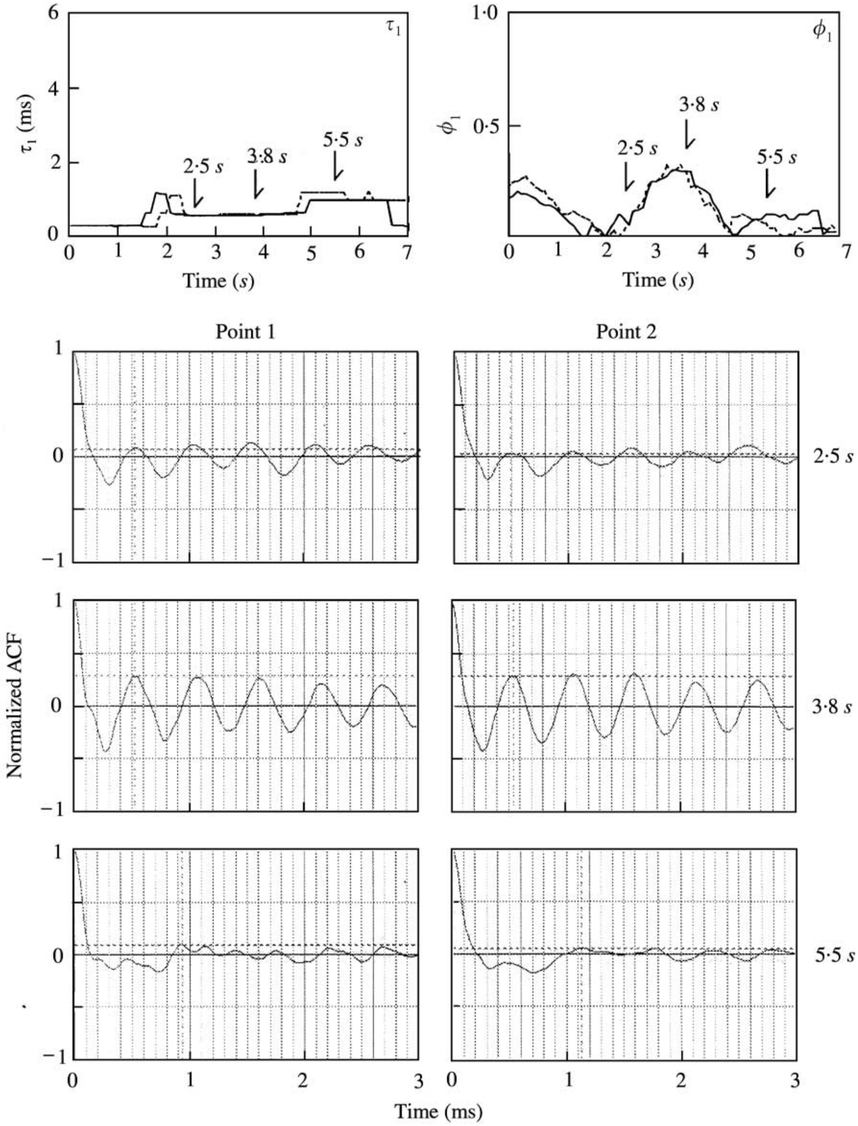

Figure 8. Waveforms of ACF for freight train for different times (2·5, 3·8, and 5·5 s) at different positions in order to indicate τ_1 and ϕ_1 activities. Upper two figures show results of running τ_1 and ϕ_1.——, Point 1; - - - - -, Point 2.

LL. Except for the noises with a tonal component, like whistling, the factors strongly related to pitch sensation (τ_1 and ϕ_1) are not significant.

The relationships between the values for four factors at Points 1 and 2 are discussed. The values for each factor were obtained at maximum $\Phi(0)$ during a single session. In addition to $\Phi(0)$ with a correlation of 0·943 (and also *LL*), τ_e has a high correlation (0·768) between the values at Points 1 and 2. The correlation value for IACC is small, 0·080, because IACC was larger at Point 2 than at Point 1 due to the reflecting surfaces. The trains were thus not perceived as a linear source.

It is quite difficult to directly compare the results of Measurement I with those of Measurement II because of the differences in train construction between countries and the

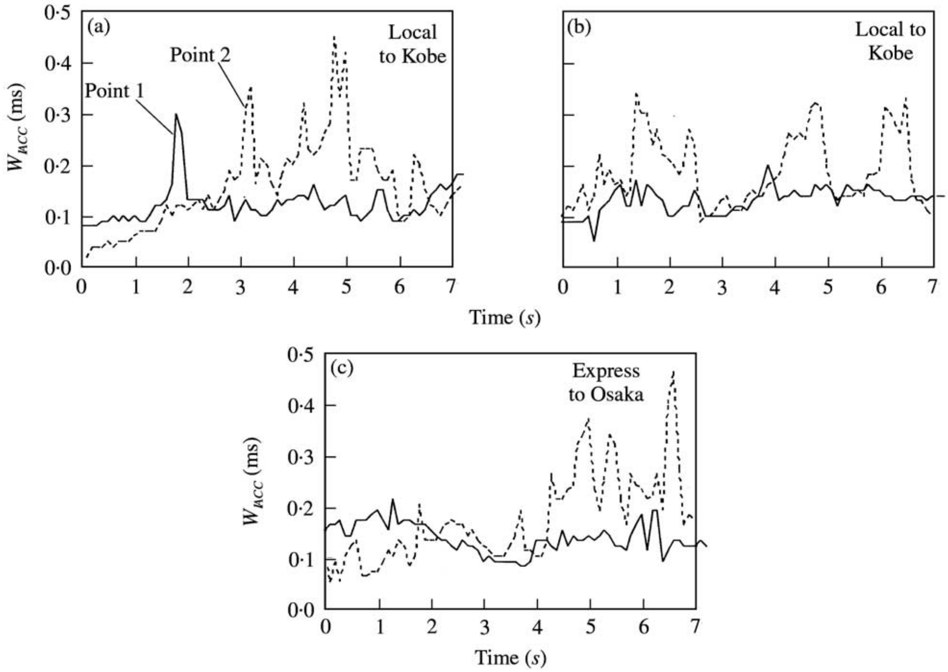

Figure 9. Typical activities of W_{IACC} for different types of trains. Straight and dotted lines represent results at Points 1 and 2 respectively: (a) result of W_{IACC} for local train to Kobe; (b) result of W_{IACC} for another local train to Kobe; (c) result of W_{IACC} for express train to Osaka.

effect of the large reflective surfaces in Measurement II. However, the same general tendencies were seen in the factors, except for IACC. The IACC of Measurement II was generally smaller than that of Measurement I because of the strong lateral reflections from the walls.

In the results for aircraft noise around an airport [6], the IACC was always near unity during the measurements, while in our results it was always less than 0·5. Another big difference was in τ_{IACC}. Aircraft noise is not a linear source, so τ_{IACC} remains almost constant, while for railway noise it changes rapidly. Thus, the difference between railway noise and aircraft noise seems to appear mainly in the spatial factors.

Applying the theory of primary sensations using the human auditory–brain model makes it possible to analyze subjective-scale values by using temporal and spatial factors obtained from environmental noise measurement, as described in section 1.

6. CONCLUDING REMARKS

By analyzing the results of measuring railway noise, we determined that different noise sources can be characterized using autocorrelation functions (ACF) and interaural crosscorrelation function (IACF) factors. For example, different noises with the same listening level *LL* can be characterized or the target noise can be identified by using other physical factors.

The proposed diagnostic system has the potential to identify the type of a noise source automatically by using the ACF and IACF factors. In our study, we used only acoustical

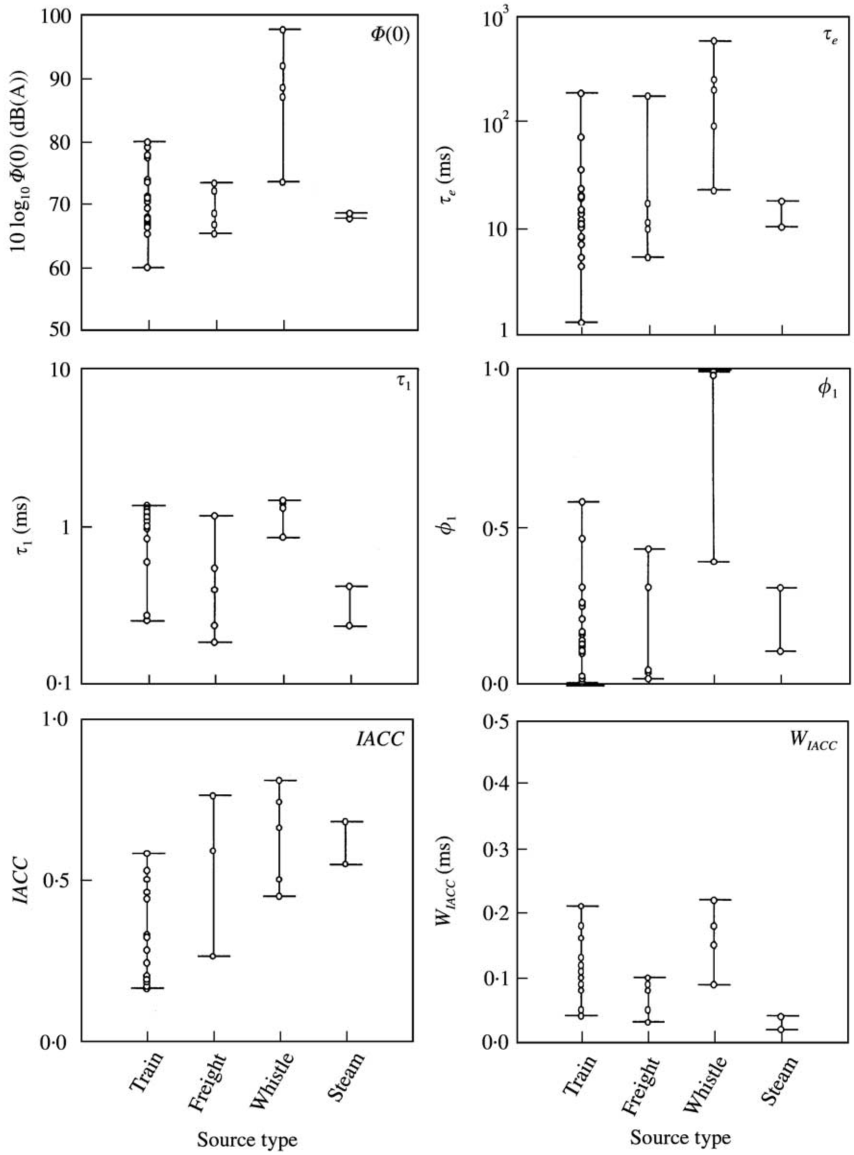

Figure 10. Distributions of values for physical factors. *LL* was excluded as its activity was the same as that for $\Phi(0)$. Error bars represent ranges of values.

TABLE 1

Estimated t-values for factors; those for "steam" were eliminated due to insufficient data

	$10 \log \Phi(0)$	$\log \tau_e$	$\log \tau_1$	$\log \phi_1$	*LL*	IACC
Train–freight	− 0·287	− 0·335	3·27†	0·223	− 0·007	0·431
Train–whistle	− 3·63†	− 3·80†	− 0·455	− 5·17‡	− 3·98†	0·228
Freight–whistle	− 4·69‡	− 3·40†	− 3·23†	− 3·91†	− 5·11‡	0·695

$^{\dagger}p < 0{\cdot}05$. $^{\ddagger}p < 0{\cdot}01$.

information to evaluate railway noise. A more specific approach would be to also use visual information captured using a video camera. Both types of information could be acquired automatically for railways by entering timetable information into the diagnostic system. We also used a dummy head with a spherical shape. To obtain spatial information in relation to any direction (up and down, front and rear), some improvement may be necessary. In our measurements, however, direction and location of the noise source were known in advance, so the dummy head did not cause any problems. To further evaluate the validity of the model, subjective experiments with humans need to be performed.

ACKNOWLEDGMENTS

We thank Toshihiro Kitamura and Ryota Shimokura for their co-operation and considerable assistance in the measurement of railway noise in Kobe. Also, we wish to thank Dr Shin-ichi Sato for his valuable discussion about the work. This work was supported by Research and Development Applying Advanced Computational Science and Technology, the Japan Science and Technology Corporation (ACT-JST), 1999.

REFERENCES

1. RESEARCH COMMITTEE OF ROAD TRAFFIC NOISE IN ACOUSTICAL SOCIETY OF JAPAN 1999 *Journal of the Acoustical Society of Japan* **55**, 281–324. ASJ Prediction Model 1998 for road traffic noise report from research committee of road traffic noise in Acoustical Society of Japan (in Japanese).
2. M. SAKURAI, H. SAKAI and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 19–27. A computational software for noise measurement and toward its identification.
3. Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 3–18. A theory of primary sensations and spatial sensations measuring environmental noise.
4. Y. ANDO 1998 *Architectural Acoustics—Blending Sound Sources, Sound Fields, and Listeners*. New York: AIP/Springer-Verlag.
5. P. FRAISSE 1982 in: *Psychology of Music* (D. Deutsch, editor). Orlando, FL: Academic Press. Rhythm and Tempo. (Chapter 6).
6. H. SAKAI, S. SATO, N. PRODI and R. POMPOLI 2001 *Journal of Sound and Vibration* **241**, 57–68. Measurement of regional environmental noise by use of a PC-based system. An application to the noise near airport "G. Marconi" in Bologna.
7. K. FUJII, Y. SOETA and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 69–78. Acoustical properties of aircraft noise measured by temporal and spatial factors.
8. Y. ANDO, S. SATO and H. SAKAI 1999 in *Computational Acoustics in Architecture* (J. J. Sendra, editor). Ashurdt Lodge: Computational Mechanics Publication, WIT Press. Fundamental subjective attributes of sound fields based on the model of auditory–brain system.
9. M. INOUE, Y. ANDO and T. TAGUTI 2001 *Journal of Sound and Vibration* **241**, 105–116. The frequency range applicable to pitch identification based upon the auto-correlation function model.
10. K. SAIFUDDIN and Y. ANDO 2001 *Journal of Sound and Vibration* **250**. Duration experience during listening to noise and tone.
11. J. Y. JEON 2001 *Journal of Sound and Vibration* **241**, 147–155. Subjective evaluation of floor impact noise based on the model of ACF/IACF.
12. Y. SOETA, K. OHTORI, K. FUJII and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 79–86. Measurement of sound transmission by hall doors.