Audio Engineering Society

Convention e-Brief 640

Presented at the 150th Convention ,Online
2021 May 25-28

A comparative perceptual evaluation of thunder synthesis techniques

Joshua D. Reiss[1], Hazar Emre Tez[2], and Rod Selfridge[3]

[1] *Queen Mary University of London, Mile End Road, London E1 4NS*

[2] *Queen Mary University of London, Mile End Road, London E1 4NS*

[3] *KTH Royal Institute of Technology, Brinellvägen 8, 114 28 Stockholm, Sweden*

Correspondence should be addressed to joshua.reiss@qmul.ac.uk

ABSTRACT

The sound of thunder is widely used in game, film and virtual reality sound design. It is also a phenomenon for which we seek a better understanding of the physics underlying the sound. Though many models of thunder have been proposed, there has not yet been a formal perceptual evaluation of the models to assess their realism and sound quality. Here, we present and evaluate the implementation of several thunder sound effect synthesis models. The models include different physical modeling and signal-based approaches, as well as a recorded sample. Evaluation was with over 50 participants. The results showed that none of the models were close to the recording in terms of realism, though signal-based models slightly outperformed the physical models. This highlights the need for comparative perceptual evaluation in sound synthesis, and identifies the limitations of current thunder simulation approaches.

1 Introduction

The gaming, film and virtual reality industries rely heavily on recorded samples for sound design. This has inherent limitations since the sound is fixed from the point of recording, leading to drawbacks such as repetition, storage, and lack of perceptually relevant controls.

Procedural audio offers a more flexible approach by allowing the parameters of a sound to be altered and sound to be generated from first principles. This reduces repetition and memory requirements, enables sound designers to achieve very specific sounds, and allows these sounds to interact with the physics of the environment [1]. However, procedural audio is not yet widely adopted in industry, partly due to current procedural audio models not sounding as realistic or as high quality as sample-based audio. By creating more realistic models and increasing the number of sounds that can be modelled, procedural audio may one day transform the industry.

A natural choice for procedural audio is environmental sounds. They occur widely in creative industries content, and are notoriously difficult to capture. On-location sounds often cannot be used due to recording issues and unwanted background sounds, yet recordings from sample libraries are rarely a good match to an environmental scene.

Thunder in particular, is highly relevant. It provides a sense of the environment and location, but can also be used to supplement the narrative and heighten the tension or foreboding in a scene. There exist a fair number of methods to simulate thunder. But to date, they all lack certain features and have had little formal evaluation of their quality. The goal of this work is to investigate, implement and evaluate existing thunder synthesis methods, in order to inform the development of a high performance, interactive thunder model.

2 Background

Lightning, shown in Figure 1, is the occurrence of a short duration, high voltage natural electrical discharge between a cloud and ground or within a cloud, accompanied by a bright flash and thunder. Lightning consists of segments, referred to as tortuosity of the lightning channel, usually between 5 and 70 meters long [2], with branching off the main path.

Figure 1. Lightning (copyright James Insogna, 2011, CC BY-NC-ND 2.0.

When the discharge happens, the temperature and air pressure in a channel of air rise rapidly to ~24000K and 10^6 Pa respectively [3]. This creates a shock wave expanding at roughly 3000 m/s [4]. Air in the channel then quickly cools down. The pressure behind the shock wave will momentarily drop below atmospheric pressure due to the inertia of the outwards traveling air mass. The shock wave will travel some distance (the *relaxation radius*), and then leave behind the *weak shock wave*. We hear these pressure waves caused by the rapidly heating air in the lightning channel. But the thunder sound is far more complex than the *clap* of a lightning strike. We hear the thunder *rumble* and *roar*, often for many seconds, as well as multiple strike sounds from different directions and at different times, and the strength of these components can depend on the distance of the listener. These aspects of the thunder sound are explained below in discussion of the thunder synthesis models.

3 Thunder synthesis

3.1 Existing thunder models

The work of Ribner & Roy [5] is perhaps the most thorough and well-known thunder model, and acted as a starting point for finding existing thunder simulations. We searched through all references they cited and all papers that have cited any of those references in order to identify any relevant simulations. For all of the papers identified that concerned thunder modelling, we then repeated the procedure, searching all citations therein and all citations of those papers. This procedure was repeated until no new potentially relevant references could be found. Potentially relevant simulations were also found based on keyword searches in databases, and search engines and the authors' prior knowledge. This exhausive procedure uncovered a large body of work on modeling, measuring and understanding lightning and thunder, and related phenomena. But only ten thunder models were found.

These models are described in Table 1. They fall into two main categories; signal-based models, which are informed by the physics but aim to reproduce the sound rather than directly model the phenomenon; and physical models which directly simulate the known physics, and the resultant sound is a natural by-product of the simulation. All the signal-based models were physically inspired, that is, the choice of processing was based on knowledge of the phenomenon, even if not directly simulating it. The physical models are either based on the work of Roy & Ribner [5] or the work of Few [4,6].

Of the ten models that were found, [7-11] are only available as software without a technical paper, and [5,12,13] provide papers describing aspects of the model but no software. Only [14] provided publicly available code accompanying a full technical description, and [15] provided software for the model described in the paper upon request.

Author	Approach	Method	Available	Platform	Real-time	Lightning	Reference
Brooks*	Signal-based	Physically-inspired	Yes	Max MSP	Yes	No	[9]
Cundangan			Yes	Max/Unity	Yes	3D	[7]
Farnell*			Yes	PureData	Yes	No	[14]
Selfridge*			Yes	JavaScript	Yes	2D	[8]
Ribner	Physical	WM Wave	No	?	?	3D	[5]
Unknown			Yes	Matlab	No	3D	[11]
Glassner			No	?	?	3D	[12]
Blanco*			On request	Matlab	No	2D	[15]
Dunkin			No	Matlab	No	3D	[13]
Saksela*		Brode waves	Yes	JavaScript	No	3D	[10]

Table 1. Known thunder simulation models, found either in the literature or online. Models with an asterisk next to the first author name were used in evaluation.

3.2WM-wave physical models

Most physical models are based on [5]. In the original paper, lightning is modeled as a 3D random walk (x ground distance, y height), originating from the origin, which is the contact point with the ground. The angle between successive segments is based on a Poisson distribution with memory dependence.

Later implementations used different approaches to construction of the lightning channel, often for ease of implementation.

WM Waves: The sound emanating from the lightning channel is based on N-waves emanating from each finite line segment along the channel, giving the shape of WM waves [16] (named for the authors of the original paper, Wright and Medendorp) arriving at the listener. Summation of these WM waves, with additional processing for effects such as attenuation with distance, filtering by air, and reverberation, gives the acoustic signal received by the listener.

We define

- r: Distance from contact point to observer.
- c: Speed of sound.
- T: Duration of WM Wave.
- l: Length of line segment, assumed to be 3m.
- ϕ: Angle between normal to segment and line connecting midpoint to observer.

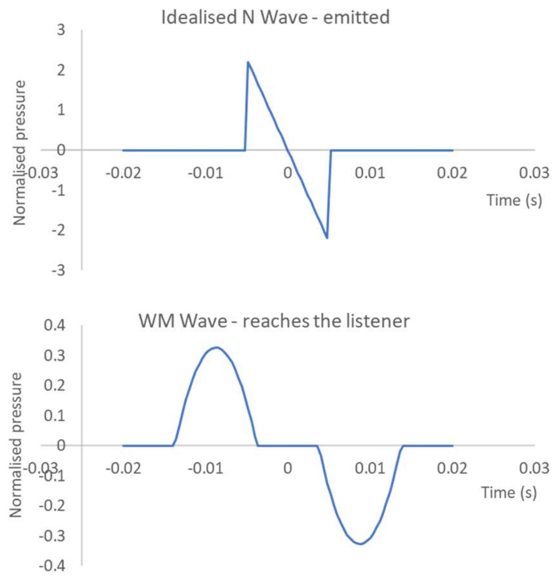

Figure 2. WM Wave for c=343 m/s, T=0.005ms, l = 3 m, B=1. On top, the emitted N wave for ϕ=0. On bottom, the WM wave received by a listener for ϕ=90°.

which yield the following dimensionless quantities;

- $\tau = (ct-r)/l$ scaled, shifted time
- $\psi = cT/l$ normalized N wave duration
- $B= l^2/(2rcT)$ amplitude coefficient

Values of T and l can vary both for line segments within a model and between models, but typical values are 5 ms and 3m respectively.

Suppose the listener is at point L and the line segment is from $\mathbf{P_1}$ to $\mathbf{P_2}$. The angle ϕ is given by Eq. 1[1];

$$\phi = \pi/2 - cos^{-1}(|\mathbf{S}\cdot\mathbf{V}|/(|\mathbf{S}||\mathbf{V}|))$$

$$\text{Where } \mathbf{S} = \mathbf{P_2} - \mathbf{P_1}, \mathbf{V} = \mathbf{L} - (\mathbf{P_2} + \mathbf{P_2})/2 \quad (1)$$

Then, for a given line segment, the WM wave can take one of two forms, as in Eq. 2 [17];

If $sin\phi < \psi$

$$-\frac{B(\psi^2 - (\tau + sin\phi)^2)}{sin\phi} \quad -\psi - sin\phi < \tau < -\psi + sin\phi$$

$$-4B\tau \quad -\psi + sin\phi < \tau < \psi - sin\phi$$

$$\frac{B(\psi^2 - (\tau - sin\phi)^2)}{sin\phi} \quad \psi - sin\phi < \tau < \psi + sin\phi$$

Else

$$\frac{B(\psi^2 - (\tau + sin\phi)^2)}{sin\phi} \quad -\psi - sin\phi < \tau < \psi - sin\phi$$

$$0 \quad \psi - sin\phi < \tau < -\psi + sin\phi$$

$$-\frac{B(\psi^2 - (\tau - sin\phi)^2)}{sin\phi} \quad -\psi + sin\phi < \tau < \psi + sin\phi$$

(2)

Thus, an N-wave of initial duration T results in a WM-wave of duration $2(l|\sin\phi|/c+T)$ when received by the observer, starting at time $r/c-T-l|\sin\phi|/c$, and ending at time $r/c+T+l|\sin\phi|/c$.

This is illustrated in Figure 2. The initial N wave is perceived as a WM wave with positive and negative portions of a sine wave, separated by a line segment.

N waves are created along the channel (every 3 m in [5]), and summed to produce the acoustic signal.

Atmospheric refraction: The flight time for each WM-wave needs to account for atmospheric refraction, which changes the length of the waves and alters their flight path from linear to curved. For N-wave number i, as the distance to observer r_i increases, the increased flight time gives new values for ψ and B when computing the WM-wave, shown in Eq. 3:

$$\psi_i = k^{1/2}\psi, B_i = kB$$

$$\text{Where } k = ln(1000/10)/\,ln(r_i/10) \quad (3)$$

So as distance increases, k decreases. Hence, amplitude and duration decrease compared to what they would have been. This effect was modeled in [12,5], but not in other WM wave models.

Multistrike: a lightning flash consists of several successive strokes, of order 60 ms apart in time, along the same tortuous channel. Thus there is a superposition of single-stroke thunder signatures with appropriate time delays. This is rendered by simply passing the summed WM waves through an FIR filter.

3.3 Brode pulse physical models

A quite different approach to physical modeling of thunder was taken in [10]. Here, lightning was approximated as straight lines with a random length of 5 to 70 meters. The lines zig-zag with a random variation of about 16 degrees between each line, and a small statistical deviation in the vertical direction. This generates segments of average 45 m length. It then adds micro-/mesotortuosity (small randomness between the macrotortuos segments). Branching of the lightning and the in-cloud crown of lightning was also simulated.

Lightning channels were generated based on the statistics given by [2]. Lightning is divided into multiple small segments, each modeled as a separate spherical sound source emitting a Brode pulse [18], shown in Figure 3, where the sharp wavefront reaches the listener first. The shape of such a pulse is given by Eq. 4.

$$y = -0.0279x - 0.376x^2 - 0.485x^3 + 0.596x^4 - 0.1259x^5 \quad (4)$$

Just as with the WM waves, this wave evolves over time. The length of the positive-pressure pulse at the ground, L, is given in Eq. 6.

$$L = R\,[0.386 + 0.147(\ln\left(\frac{H}{10.46R\cos\theta}\right) - \frac{H}{16{,}000})]^{2/3} \quad (5)$$

Where H is the height of the pulse source in meters, R is the initial radius of the pulse (typically 5 m) and θ is the angle between the vertical and the listener.

[1] math.stackexchange.com/questions/472008/angle-between-line-and-a-plane

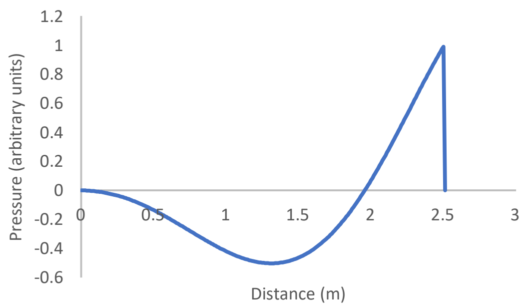

Figure 3. Brode pulse.

To generate sound, each segment is broken into many (10s to 1000s) smaller segments, each one producing the pulse mentioned above. All of these pulses are summed together to produce the output signal.

The steps in the model are as follows;

- Establish the main lightning channel path, branches and crown
- Create an array of vertices on the path. There are about 200 of these. Each is a 3d point, z is height and the *x* axis is the line between listener and where the strike hits
- Now sample the lightning channel as a large collection (hundreds of thousands) of points
- For each point
  - Determine when it reaches the listener
  - Create a pulse of a given duration
  - Apply the pulse at the correct time
- Sum the pulses to produce the sound
- Add reverb
- Add attenuation

For the last two steps, distance-based filtering and simulated reverb (convolution with an exponentially decaying random sequence) were applied.

3.4 Signal-based procedural models

In [14], Farnell introduced a procedural thunder model. It uses knowledge of the physics and perception of thunder to guide implementation, but does not directly implement any physics. Key elements include: strike pattern generator (a sequence of values that get successively smaller and further separated in time); single strike sound generator (shaped, filtered burst of noise); multi-strike sound (applying the strike pattern generator to several single strike sounds); damped N-wave rumble generator (a randomly rising and falling ramp which is then shaped to produce parabolic pulses with a controllable density, aiming to mimic the rumble of N waves); deep noise, based on distorted, low-pass filtered noise to give a low frequency rumble ; after-image (reflections from the environment) and environmental echoes.

The model of [14] was expanded upon and re-implemented in the FXive system [8], now available as Nemisindo [2], an online hub that stores many procedural sound effect synthesis engines designed to be used from the browser. This model mostly replicates the strike-pattern from [14], but with accompanying graphical rendering of lightning. Randomisation was also used to mimic the tortuosity of a lightning channel, and other minor changes to port the code to conform with the Web Audio API, including simplification of the distance-based filtering to avoid clipping.

In [9], the initial strike was generated as a series of short, staccato noise envelopes, to emulate the superposition of N-waves heard directly from the lightning. A series of very close together bangs are produced, which imitates the in-phase N-waves of a tortuous channel. Hundreds of pulses are triggered using delay lines. These pulses are staggered by changing the delay in relative ratios every time a thunder is being generated.

The impulses are used to create amplitude envelopes to shape noise. 110 *N*-waves are generated from noise with amplitude envelopes based on observations of sonic booms made by NASA (personal communication from the author of [9]). A resonant lowpass filter and short delays are then used to emphasize the low end of noise strikes and give more randomization to the initial strike. The filter simulates sound absorption by air. Finally, distortion is used once elements of the after image were blended into the signal with a wave-shaping function to flatten the high-end of the strike and give it a more clap-sound.

To create an afterimage, the noise envelopes are further filtered using different cutoff frequencies and

[2] https://nemisindo.com

delay values to simulate close (1.25s), medium range (2.5s) and long range (4.5s) reflections. After this, the envelopes are processed using multiple feedback loops and filters to add trailing to the sound. The filtered and amplitude attenuated copies of the signal are then fed back to further reflections in a loop that simulates thunder bouncing off very distant surfaces.

Finally, [7] was just a simplification of [9], removing elements that the author felt were unnecessary, and then reimplementing this in Unity.

4 Thunder simulation evaluation

Subjective evaluation of thunder synthesis was based on the approach in [19], whereby different synthesis models were evaluated against each other and a recorded sample (the reference) in order to determine which synthesis method produces the most realistic result.

Evaluation was performed using the Web Audio Evaluation Toolkit [20,21], which provides a platform for perceptual audio evaluation experiments. Evaluation included samples from the five highlighted synthesis techniques from Table 1; Blanco [15], Selfridge [8], Farnell [14], Brooks [9] and Saksela [10]. A recorded sample of real thunder was also used as a reference, downloaded from the BBC sound effects archive.

The audio perceptual evaluation (APE) method [22] was applied. This is a multistimulus paradigm to present a user with a continuous scale (Very unrealistic - Quite unrealistic - Quite realistic - Very realistic) where samples can be played and dragged across the scale to rate them. Whether the participant had audio experience was confirmed in order to compare how the model performs for audio professionals and inexperienced participants. For consistency, all samples were set to the same loudness and a 44.1 kHz sample rate. For each participant, stimuli were presented in randomized order.

51 volunteer participants were each asked to rate a set of samples in terms of realism. All participants reported normal hearing and all had experience with audio by either playing an instrument or working with audio-related technology. Participants did the test remotely during the Corona virus lockdown, so under varying listening conditions.

5 Results

Figure 4 gives the results of the multi-stimulus test, including mean rating and 95% confidence intervals for perceived realism of each sample. It is clear that none of the models produce a result that comes close to matching the recorded sample in terms of realism.

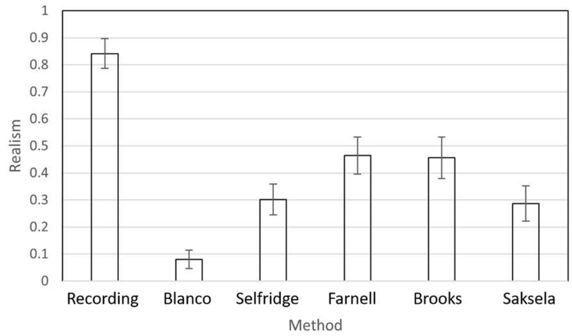

Figure 4. Subjective evaluation results for five thunder synthesis models and a real recording. Error bars are for 95% confidence intervals.

6 Conclusions

This paper highlights a previously unknown issue. The prior art emphasises the quality of their techniques, but to the best of our knowledge, this is the first time that those techniques were formally and subjectively compared against each other and against thunder recordings. None of the techniques presented a high level of realism. This may be in part due to the particular settings that were applied to generate thunder, and for one technique [15], the provided source code may not be the same version that was used in the paper. Yet it is clear that the listening test gave no evidence that any of the proposed simulation techniques was sufficiently realistic.

Interestingly, the Farnell method performed significantly better than the closely related FXive method, though this could simply be due to the choice of samples from each method. Blanco's method, the only one that tries to closely reproduce [5], performed worst. However, it is highly unlikely that the code we received from the author was the final version that was used in [15].

7 Future research directions

Implementation and investigation of all the methods highlighted some approaches to building a better model.

- for the lightning channel, include the crown of electrical discharges within the cloud, and the branches off from the main branch
- apply multistrike generation
- apply delay, filter and reverb carefully to the generated thunder sound sources. Testing of implementations showed that reverb in particular had a huge effect on perception.
- represent the geometry of the lightning channel in three dimensions
- distance attenuation should be frequency dependent
- include natural reverberation of an outdoor environment
- the implementation should be stereo or immersive sound
- include recent measurements of thunder features [23-25]
- incorporate an HRTF if listened to over headphones

It should be noted that all models produced monaural output, a major limitation. None of the models attempted convolution with a recorded impulse response for reverberation.

Filtering by air was also poorly approximated. This frequency-dependent can be calculated according to ISO 9613-1:1993 [26]. It should cause distant sounds to "rumble" while the closest sounds are sharp and discernible. Assuming fixed temperature, pressure and humidity, one should model sound absorption from air as a function of distance by using a low pass filter whose cut-off frequency is proportional to the reciprocal of distance, similar to the method described in [27].

Finally, we note that while the samples used for evaluation were based on unmodified versions of the thunder simulation models, we attempted modifying and improving the models. This is still a work in progress, but its clear that major computational improvements can be achieved. Initial testing has suggested that the physical models, even with enhancements, should be able to operate real-time.

Additional Resources

Data, code and sound samples are available at https://github.com/joshreiss/thunder-simulation-evaluation

Acknowledgements

We would like to thank the authors of all the thunder simulation models, especially those who made the models publicly available (Andy Farnell, Joe Brooks and Kai Saksela) or provided code on request (Francesco Riggi).

References

[1] R. Selfridge, D. Moffat, E. J. Avital, J. D. Reiss, “Creating real-time aeroacoustic sound effects using physically informed models,” Journal of the Audio Engineering Society, vol. 66, no. 7/8 (2018).

[2] V. A. Rakov, M. A. Uman, Lightning: physics and effects (Cambridge University Press) (2003).

[3] R. E. Orville, “A high-speed time-resolved spectroscopic study of the lightning return stroke: Parts I-III,” Journal of the Atmospheric Sciences, vol. 25, no. 5, pp. 827–856 (1968).

[4] A. A. Few, “Acoustic radiations from lightning,” Handbook of atmospheric electrodynamics, vol. 2, pp. 1–31 (1995).

[5] H. S. Ribner, D. Roy, “Acoustics of thunder: A quasilinear model for tortuous lightning,” Journal of the Acoustical Society of America, vol. 72, no. 6, pp. 1911– 1925 (1982).

[6] A. A. Few, “Acoustic radiations from thunderstorms,” in N. R. Council, G. S. Committee, et al. (Eds.), The Earth’s electrical environment (National Academies Press) (1986).

[7] O. Cundangan, “Thunder Synthesis in Unity,” http://medium.com/@othnielcundangan/mumt-307-project-thunder-synthesis-in-unity-5274643f22e7 (2017).

[8] P. Bahadoran, A. Benito, J. D. Reiss, E. Avital, “FXive: AWeb Platform for Procedural Sound Synthesis,” presented at the 144th Audio Engineering Society Convention (2018). Now available at https://nemisindo.com

[9] J. G. Brooks, “Thunder generator for MaxMSP,” https://joebrookssound.com/2016/11/30/thunder-generator-for-maxmsp/ (2016).

[10] K. Saksela, “Thunder simulation,” https:// github.com/kai5z/Thunder-simulation, https://blog.kaistale.com/?p=1340 (2014 Sept.).

[11] Anonymous, ‘Thunder Generator’, https://forum.pdpatchrepo.info/topic/1028/synthetic-thunder (2007)

[12] A. Glassner, “The digital ceraunoscope: synthetic thunder and lightning. 2,” IEEE computer graphics and applications, vol. 20, no. 3, pp. 92–96 (2000).

[13] J. Dunkin, D. Fleisch, “Digital Simulation of Thunder from Three-Dimensional Lightning,” APS, vol. 55, no. 4, pp. P1–032 (2010).

[14] A. Farnell, Designing sound (MIT Press Cambridge, UK) (2010).

[15] F. Blanco, P. La Rocca, C. Petta, F. Riggi, “Modelling digital thunder,” European Journal of Physics, vol. 30, no. 1, p. 139 (2008).

[16] W. M. Wright, N. W. Medendorp, “Acoustic Radiation from a Finite Line Source with N-Wave Excitation,” The Journal of the Acoustical Society of America, vol. 43, no. 5, pp. 966–971 (1968).

[17] D. Roy, "A Monte Carlo model of tortuous lightning and the generation of thunder," Thunder," Univ. Toronto Inst. Aerospace Studies Rep. 243 (CN ISSN 0082-5255 (1981): 2251-2251.

[18] H. L. Brode, “Numerical solutions of spherical blast waves,” Journal of Applied physics, vol. 26, no. 6, pp. 766– 775 (1955).

[19] D. Moffat, J. D. Reiss, “Perceptual Evaluation of Synthesized Sound Effects,” ACM Transactions on Applied Perception (TAP), vol. 15, no. 2, p. 19 (2018 March).

[20] N. Jillings, B. De Man, D. Moffat, J. D. Reiss, “Web Audio Evaluation Tool: A Browser-Based Listening Test Environment,” presented at the Proc. Sound and Music Computing 2015 (2015 July).

[21] N. Jillings, B. De Man, D. Moffat, J. D. Reiss, “Web Audio Evaluation Tool: A framework for subjective assessment of audio,” presented at the Proc. 2nd Web Audio Conference (2016 April).

[22] B. De Man, J. D. Reiss, “APE: Audio Perceptual Evaluation Toolbox for MATLAB,” presented at the Audio Engineering Society Convention 136 (2014).

[23] J. Wang, et al. "Characteristics of acoustic response from simulated impulsive lightning current discharge." High Voltage 4.3 (2019): 221-227.

[24] A. Lacroix, et al. "Acoustical energy of return strokes: A comparison between a statistical model and measurements." Geophysical Research Letters 46.20 (2019): 11479-11489.

[25], J. A. P. Bodhika, "A brief study on thunder claps." Applied Acoustics 145 (2019): 98-103.

[26] Intl. Standards Organisation, ‘ISO 9613-1:1993 Acoustics — Attenuation of sound during propagation outdoors — Part 1: Calculation of the absorption of sound by the atmosphere’, 1993

[27] J. Huopaniemi, L. Savioja, M. Karjalainen, “Modeling of reflections and air absorption in acoustical spaces a digital filter design approach,” Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 4–pp (1997).