12th INTERNATIONAL SCIENTIFIC CONFERENCE ON DEFENSIVE TECHNOLOGIES
OTEH 2026
Serbia, Tara, 21 - 24 October 2026

# IMPROVING DRONE DETECTION USING ACOUSTIC METHODS AND STATISTICAL FILTERING

MIODRAG VRAČAR
Military Technical Institute, Beograd, vracarmiodrag@mts.rs

STEVO VRAČAR
ITS Information Technology School, Beograd, stevo5817@its.edu.rs

RADIANA BRUSIN
Military Technical Institute, Beograd, radiana.brusin@mod.gov.rs

***Abstract:*** *Acoustic methods for the detection of propeller-driven drones are used at relatively short distances, up to a few hundred meters. The drones are relatively weak sources of sound. The appearance of acoustic disturbances in the environment significantly reduces quality of the detection and localization drones by acoustic methods. Starting assumption, in resolving of this problem, is that acoustic ambient noise has a Gaussian natural distribution. A significant reduction in the influence of ambient noise was achieved by using the so-called statistical filter, which eliminates all contributions of ambient noise if their distribution is Gaussian. The function of filtering acoustic signal, that is, the filter itself, is realized as an independent module, that is, a processor that analyzes the acoustic signal using standard statistical methods, as well as higher order statistics. To check the quality of the filter itself it is used the five core quality factors: correlation coefficient, Normalized Root Mean Square Error (NRMSE), Signal-to-Error Ratio (SER), Maximum Absolute Error (MaxAE) and Estimated Delay (τ). The analyzed acoustic signals were generated in different flight mode of the drone, such as hovering, circular movement, left-right movement and others.*

***Keywords:*** *drone, acoustic noise, statistical filter.*

## 1. INTRODUCTION

Acoustic detection of drones is crucial becausee it fills critical gaps left by radar and cameras against modern drone threats, especially in complex environments. There are more significant reasons for using acoustic methods such as its passive and steatlthy nature, identification of fiber optic drones, and in addition, these acoustic methods can act as a low-power trigger for turnig on radar and high resoluton cameras only after it hears a drone nearby.

There are different ways to do make signal useful and clear, such as filters and transforms. Nowadays exists and more advanced techniques like computer programs and smart networks. This paper explores a relatively little-used method for extracting useful signals from noise, which is statistical method based on higher order statistics. In adddition to the description of the filter itself, the paper presents its application in acoustics, seismics, hydroacoustics and infrasound analysis. Several tests were performed to obtain an evaluation of quality of the filter. The five core quality factors are used: correlation coefficient, Normalized Root Mean Square Error (NRMSE), Signal-to-Error Ratio (SER), Maximum Absolute Error (MaxAE) and Estimated Delay (τ). All of these methods help us to understand how useful is this filtering method for extracting a signal from noise.

Statistical filters use well-known statistical quantities, like mean and variance to process the signal. The mean filter, for example, replaces each data point with the average of neighboring points within a specific window. By averaging nearby values, high-frequency noise is smoothed out, and the signal is effectively de-noised, see [1]. Practical explanation of moving average filters with implementation examples and applications are given in the paper Smith, Steven W. [2]. Moving average is effective at removing high-frequency noise, but may blur sharp features or sudden changes in the signal. The choice of the window size is critical; a too small window might not remove enough noise, while a too large window can lead to loss of important signal details. There are many others filtering techniques. The Gaussian smoothing involves convolving the signal with a Gaussian kernel. The Gaussian function gives more weight to central points and less weight to distant points, see [3, 4].

Wavelet technique is used for de-noising signals by thresholding the wavelet coefficients. Wavelet transform decompose the signal and transform it into wavelet coefficients at different scales and positions. These coefficients represent the signal's details at various levels of time and frequency resolution [5, 6].

Classic comprehensive treatment of statistical filtering theory is well described in the book of the author Jazwinski, Andrew H. [7]. Modern Bayesian approaches including particle filters is described by Särkkä S. [8]. Comprehensive modern treatment of statistical filters including Kalman filters [9], particle filters, and their

**Corresponding author:** M. Vračar;
E-mail: vracarmiodrag@mts.rs
DOI: 10.5937/OTEH26054V

applications in tracking and estimation is described too. Particle filtering is based on sequential Monte Carlo methods for nonlinear/non-Gaussian problems such as target tracking, navigation, signal processing, finance, robotics, and others.

This paper presents the results of filtering acoustic signals, drone noise, using a statistical filter whose working principle is completely based on statistical methods. In the implementation of the filter, in addition to classical statistical procedures, such as the calculation of mean values, the so-called higher order statistics, i.e. the fourth order cumulants are used too. All the mentioned procedures are applied on signal segments – windows, with or without overlapping.

## 2 PRINCPLES OF THE FILTER DESIGN

Using the fourth-order cumulants (also called kurtosis in the statistical world) is a game-changer for machine noise analysis because they are "blind" to Gaussian noise and excel at detecting non-linear, impulsive signatures – exactly what failing or, standard mode of work, machine produce. Using them enables to:

- Suppress background noise (real-world acoustic sensors pick up wind, traffic, and electronic hiss-all Gaussian (random) noise. Fourth-order cumulants theoretically become zero for Gaussian noise, meaning they filter out the environment and leave only the machine unique "fingerprint".
- Reveals non-linear interactions. Machine parts, or drone propellers, interact in complex ways. The fourth-order cumulants capture phase coupling between frequencies, showing their interconnections, not just that they exist.
- Immune to synchronous noise. If propeller creates a loud, constant tone, the second-order methods see it as a big peak. But cumulants suppress these steady periodic noises, letting us to focus on the changing signatures.
- The method is robust to signal distortions. Mathematically it remains stable even if the microphone is cheap, or the signal clips slightly.

Relatively bad argument for using the fourth-order cumulants is that they require a lot of data (longer sampling time) to estimate accurately, and are computationally heavy. But for critical machinery – like drones, the early detection of its presence is worth the cost.

The flow chart diagram of the filter is shown in Figure 1. The process of statistical filtering is carried out in several steps. The first is segmentation with or without overlap. This is a significant considering that in order to successfully calculate the fourth-order cumulants, dynamically in time, a larger amount of data is required. The fourth-order cumulants are calculated over thus formed data segments, which contain information about the signatures of interest. Thus, this part of the analysis is responsible for extracting the frequencies of interest. In this way, the time-varying part of the signal was performed.

In practical applications DC signal component is also of interest. In order to preserve this component of the signal, a block named "Overall statistics of the signal segments" is created, see Fig. 1. Its basic role is to calculate the mean value of each signal segment. By adding these two branches, we get a filtered signal, or statistically filtered signal, which has information about both the DC component of the signal and all frequencies of interest of the signal.

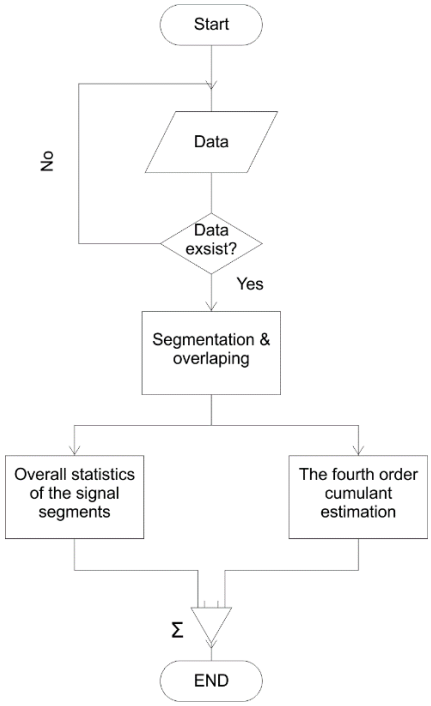



**Figure 1.** Flow chart diagram of the statistical filter algorithm.

The description of the statistical filter is given in the form of a block diagram, since its mathematical principles have been well known for a long time, and there is extensive literature on it. These signal analyses are very similar to the one who applies Short Time Fourier Transform (STFT), so it can be said that this approach opens up statistical domain in signal analysis. Statistical domain is just as equal the frequency domain considering that it closely connects them.

## 3. THE METRICS USED TO EVALUATE THE QUALITY OF THE FILTER

To check the quality of the filter itself it is used the five core quality factors: correlation coefficient, Normalized Root Mean Square Error (NRMSE), Signal-to-Error Ratio (SER), Maximum Absolute Error (MaxAE) and Estimated Delay (τ).

### 3.1 The Pearson correlation coefficient, ρ

The Pearson correlation coefficient, $\rho$, measures the linear relationship between source ($X$) and filtered ($Y$) signals:

$$\rho_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \cdot \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (1)$$

where $n$ is number of the samples.

### 3.2 Normalized Root Mean Square Error (NRMSE)

Normalized Root Mean Square Error (NRMSE) is defined by following equation:

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2}}{X_{\max} - X_{\min}}. \quad (2)$$

### 3.3 Signal-to-Error Ratio (SER)

Signal-to-Error Ratio (SER) measures how accurately the filter reproduces the desired frequency response.

$$SER = \frac{Power\ of\ the\ output\ signal\ that\ passes\ correctly}{Power\ of\ the\ signal\ that\ is\ missshape} \quad (3)$$

The "error" here isn't noise or distortion from external sources – it is the deviation of the real filter from the ideal mathematical model.

A high SER means the filter behaves almost exactly as deigned. A low SER means the filter introduces significant amplitude/phase errors.

If $H_{ideal}(f)$ = desired (ideal) frequency response, and $H_{real}(f)$ = actual measured frequency response, then the error in the frequency domain is:

$$E(f) = H_{real}(f) - H_{ideal}(f), \quad (3a)$$

and SER is:

$$SER = \frac{\int |H_{ideal}(f)|^2 df}{\int |H_{real}(f) - H_{ideal}(f)|^2 df} \quad (3b)$$

In decibels (most common):

$$SER_{dB} = 10 \cdot log_{10}\left(\frac{\int |H_{ideal}|^2}{\int |E|^2}\right) \quad (3c)$$

The main factors which contributes error in filters are: passband ripple, insufficient stopband attenuation, nonlinear phase, coefficient quantization (digital filters), and component tolerances (analog filters).

In case when we have input $(X)$ and output $(Y)$ amplitudes directly, the gain $(G)$ of a system is defined as, $Y_{ideal\ output} = G \cdot X_{input}$, the error is the difference between the actual output and the ideal expected output. Power of the signal that is misshape is estimated starting from error amplitude, $A_{error}$,, which is defined as follows:

$$A_{error} = Y_{actual\ output} - Y_{ideal\ output} \quad (3d)$$

If we don't know $Y_{ideal_output}$, $SER_{dB}$, is estimated, starting from Eq. (3), by following equation:

$$SER_{dB} = 10 \cdot log_{10} \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}. \quad (3e)$$

This definition of $SER_{dB}$ is used in this paper.

### 3.4 Maximum Absolute Error (MaxAE)

Maximum Absolute Error (MaxAE) is estimated based on the following considerations:

$$max|X_{true} - X_{estimated}|$$

$$\text{or } \mathrm{MaxAE} = \max_{i=1,\ldots,n} |Y_i - X_i|. \quad (4)$$

MaxAE is the single largest deviation between the filter's actual response and the ideal response, measured across all frequencies of interest. The significance of this error is in fact that many standards specify a maximum allowed error, not an average.

### 3.5 Estimated Delay (τ)

The delay is estimated by finding the lag that maximizes cross-correlation:

$$\hat{\tau} = \arg\max_k |R_{XY}[k]| \quad (5)$$

where cross-correlation, $R_{XY}$, is estimated by

$$R_{XY}[k] = \sum_{i=1}^{n} X[i] \cdot Y[i+k], \quad (5a)$$

for $lag\ (k)$ (with appropriate zero-padding).

### 3.6 Signal power

The total signal power is the integral (or sum) of the squared magnitude of the frequency response:

$$P_{signal} = \int_{-\infty}^{\infty} |H(f)|^2 df \quad (6)$$

Power in decibels (dB):

$$P_{dB} = 10 \cdot log_{10}\left(\frac{P}{P_{ref}}\right) \quad (6a)$$

Common references: dBm (reference = 1 mW), dBW (reference = 1 W), and dB (no reference, just a ratio).

## 4. RESULTS OF THE ANALYSIS OF THE ACOUSTIC NOISE OF DRONES IN HOVERING MODE

The quality and capabilities of the acoustic filter were checked through the analysis of the acoustic noise of five types of drones that were in hovering mode. The results of the analysis are presented graphically, showing the real signal, the filtered signal, as well as the spectrogram of the original signal.

#### 4.1.1 Drone DJI Matrice 210

DJI 210 Matrice 210 quadcopter is designed for industrial applications. Its total weight with batteries is 4.8 kg and overall dimensions 883 × 398 × 886 mm.

#### 4.1.2 Drone DJI Mavic Pro

DJI Mavic Pro drone has total mass, when include gimbal cover, of 743 g, and folded dimensions 83 × 83 × 198 mm.
Maximal speed is 65 km/h.

#### 4.1.3 Drone Parrot Anafi

Parrot Anafi is relatively small drone with overall

dimensions 244 × 67 × 65 mm and weight 320 g. Maximum horizontal speed is 15 m/s, and vertical speed 4 m/s.

### 4.1.4 Drone Yuneec Typhoon

Yuneec typhoon drone is hex copter (6 – rotors) designed for serial photography. Its weight is around 1695 g, and with maximum speed 13.5 m/s.

### 4.1.5 Drone DJI Phantom 4 Pro Plus

Total weight of this drone is 1388 g (battery and propellers are included). Maximal speed is 72 km/h, and diagonal size (excluding propellers) is 350 mm.

### 4.2 The quality of the filtrated data

It is possible to conclude subjectively about the quality of the performed filtering of the acoustic noise of drones by looking at the diagrams that show the source and filtered data, see Fig. 2. as their spectrograms, see Fig. 3, as well as objectively using parameters described in Section 3 of this paper. Both approaches have advantages and disadvanteges. The advantage of the first method is that it is fast and gives the first assessment of the acceptability of the solution. As for the second – analytical method, it is more objective because it is based on several statistical parameters.

Lookking at the graphs in Fig. 2, it can be concluded that the amplitudes of the filtered acoustic noisesignal are significantly smaller than the amplitudes of the raw data. This is consequence of the existence of a significant acousic noise that statistically has a Gaussian distribution.

Superior supression of Gaussian noise is performed, and therfore better SNR. The output have a significantly better SNR compared to linear filters (like FFT – based bandpass. The phase information is restored (better fidelity). Standard power – spectrum (second order) filters destroy phase information – they point out only what frequencies exist, but not when they occur. The fourth order cumulants preserve phase coupling. This means the output waveform preserves the shape of the impulses. The output is temporally accurate, which is critical for localizing drone by using sounds.

A 4th-order cumulant filter can isolate one source because it exploits statistical independence. The output is „cleaner“ in the sence that it belongs to only one physical source.

Linear filters struggle with „pink“ or „brown“ noise (common in industrial halls). The 4th-order filter outputs a signal free from this color, meaning the background sounds flat and uniform, making the fault signatures pop out visually and audibly.

If the acoustic noise itself is non-Gaussian (e.g., another drone or a hammering sound), the filter will not supress it. The output will be mix both sources.

In addition, because 4th-order cumulants amplify tails (large amplitudes), a single accidental loud bang (e.g., far and near explosions) can dominate the output temporarily ruining the signal quality. This unwanted propetry of the 4th-order cumulants can be overcome brcause the drone detection algorithm will omit these segments, bearing in mind that the drone represents a continuous source acustic noise.

In Table 1, the results of the statistical analysis of the acoustic signals of drones are systematized. Inspired by the multi-metric evaluation approaches found in literature [10], a novel aggregate Quality Factor (QF), see Eq. 6, defined as the weighted harmonic mean of individual normalized *RMSE*, *SER*, *MAE* ad estimated delay into a single percentage value, where the weights, $w_i$ reflect the relative importance of each metric for the target application is proposed. In this case, the weighting factors were chosen as follows: $w$ = [0.5, 0.2, 0.2, 0.1]. The greatest importance is given to the correlation factor, which is weighted by 0.5.

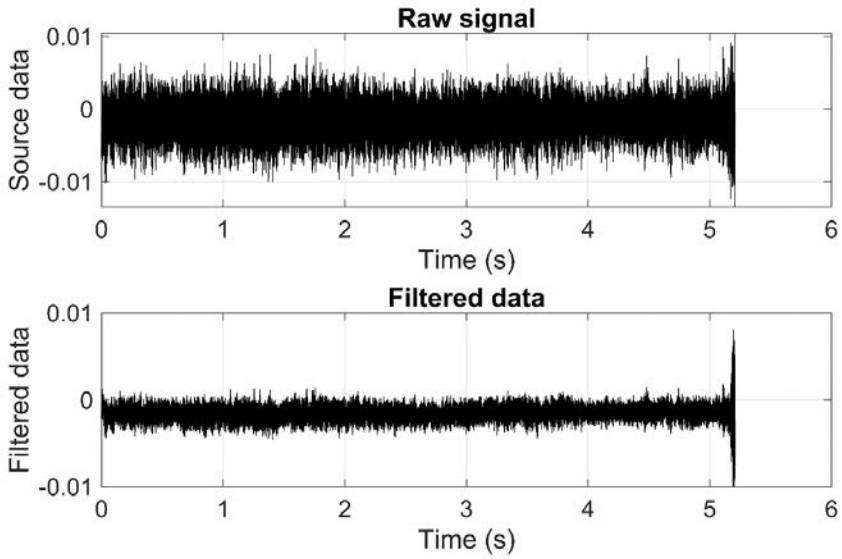



**Figure 2.** Raw and filtered data of the acoustic noise of DJI Matrice 210 drone in hovering mode.

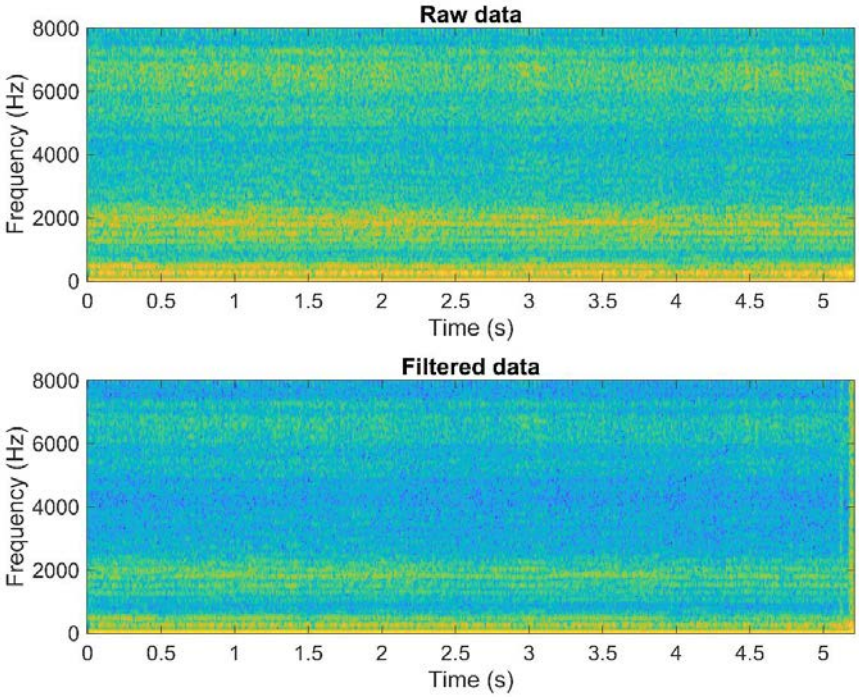



**Figure 3.** Spectrograms of raw and filtered data of acoustic noise of DJI Matrice 210 drone in hovering mode.

**Table 1.** Statistical parameters of the drone acoustic noise

| +Statistical parameter that characterize source and filtered data | Type of drone | | | | |
|---|---|---|---|---|---|
| | DJI Matrice 210 | DJI Mavic Pro | Parrot Anafi | Phantom 4 Pro plus | Yuneec Typhoon |
| Correlation, ρ | 0.9073 | 0.7248 | 0.8110 | 0.6960 | 0.8439 |
| Error_mean | -0.0039 | -0.0046 | -0.0041 | -0.0040 | -0.0039 |
| Error_std | 0.0087 | 0.0068 | 0.0053 | 0.0069 | 0.0054 |
| NRMSE | 0.0378 | 0.0837 | 0.0807 | 0.0711 | 0.0574 |
| SER_dB | 13.4172 | 15.0861 | 16.0887 | 14.3544 | 15.7793 |
| Max_Abs_Error (MaxAE) | 0.1204 | 0.0619 | 0.0663 | 0.0595 | 0.0713 |
| Delay_samples | 27 | 33 | 30 | 45 | 46 |
| Delay_seconds | 0.0032 | 0.0040 | 0.0036 | 0.0054 | 0.0055 |
| Signal_power | 0.0020 | 0.0022 | 0.0018 | 0.0017 | 0.0017 |
| Error_power | 8.9689e-05 | 6.6790e-05 | 4.4406e-05 | 6.3199e-05 | 4.4115e-05 |
| RMSE | 0.0095 | 0.0082 | 0.0067 | 0.0079 | 0.0066 |
| MAE | 0.0070 | 0.0064 | 0.0052 | 0.0063 | 0.0052 |
| Normalized_max_error | 0.4798 | 0.6344 | 0.8034 | 0.5324 | 0.6158 |
| Quality_score (%) (QF) | 64.6052 | 55.2098 | 57.8653 | 57.0072 | 62.7528 |
| Sampling_frequency (Hz) | 8333 | 8333 | 8333 | 8333 | 8333 |

$$QF = w_1 \cdot \rho + w_2 \cdot (1 - \text{NRMSE}) + w_3 \cdot \frac{\text{SER}}{60} + w_4 \cdot \left(1 - \frac{\text{MaxAE}}{0.1}\right) \quad (7)$$

By removing the Gaussian noise from the drone's acoustic noise, the result is that the spectrogram of the filtered signal, see Fig. 3, has become visually clearer, and the spectral components that originate from the operation of the drone's electric motors and its propellers are only clearly highlighted and dominate. The Gaussian noise (color disappeared in noisy areas) while preserving non-Gaussian signals (like machinery vibrations, or coms). The spectrogram's time-frequency resolution is unchanged, but the amplitude/color scale now represents non-Gaussian, not absolute power. Because 4$^{th}$-order

filters are phase-blind, the spectrogram does not show phase coupling (bispectral peaks vanish).

As it well known the spectrogram is only interpretable if the signal is stationary over the window length; otherwise, time-varying cumulants distort the color scale. Therefore it is necessary to show the original spectrogram side-by-side, see Fig. 3, otherwise, the filtered version is meaningless without a baseline.

Finally, the 4$^{th}$-order filter suppresses Gaussian noise but amplifies any non-Gaussian interference (e.g., clicks or transients), so it is necessary to inspect the spectrogram for artificial harmonic streaks caused by outliers, not real signals. These facts regarding filters based on 4$^{th}$-order cumulants point to the need to integrate this filtering principle into algorithms that will identify these outliers in the signal end reduce their impact.

## 5. DISCUSION

It is well known that for acoustic based detection technology of the drones distinguishing is the distinctive features of acoustic patterns from ambient noise, and other acoustic sources. The noise spectrum consists of line spectra and their interconnecting components, where harmonic frequencies appear as distinct spectral lines. The primary goal is that acoustic signatures of drones become clearly differentiated from interference signals and ambient noise. Additional difficulty of the acoustic drone signature investigation is scenario un-homogeneity and ambient noise variability.

According obtained results, which are generated by application of statistical filtration method, which is based on different statistical procedures, such as classic statistical parameters and using higher order statistical methods on overlapped windows of the signal, relatively good results are performed. In all cases QF factor is greater from 50%, see Table 1. Signal to noise is significantly improved. Spectral lines that belongs to the drone activity are preserved as well. The raw acoustic data signal is converted into more appropriate form for further analytical purposes.

Implementation of the statistical filtration method in acoustic drone detection algorithms is significant preprocessing step, dedicated to extracting relevant features and reducing noise ambient interference.

Ambient noise compromise drone detection capabilities. Undoubtedly denoising plays an important role in drone acoustic detection. There are other approaches to reducing the impact of ambient noise such as methodology based on the linear shrinkage – subsequence projection – power spectral density filter (LSP) that employs covariance matrices and enhanced Wiener filtering, achieving effective noise suppression within a 135 m radius, see [11].

## 6. CONCLUSION

The paper shows the application of a statistical filter in the pre-processing of the raw acoustic signals, which originate from five types of drones in the hover mode. A significant increase in the signal-to-noise ratio was achieved, while retaining all the spectral components which originate from drone's activities.

The proposed statistical method of signal filtering largely eliminates the Gaussian noise, which makes up a significant part of the acoustic signal, since higher order statistics, i.e. cumulants of the fourth order, are implemented in the filter itself.

Nowadays deep learning techniques performed substantial advancements in drone acoustic detection. That deep learning architectures includes convolutional neural networks, recurrent neural networks, convolutional recurrent neural networks, and others. Used method for filtration of the raw acoustic noise drone's data can well serve to create better database for training advanced neural networks.

## References


[1] OPPENHEIM, ALAN V. AND SCHAFER, RONALD W. AND BUCK, JOHN R.: *Discrete-Time Signal Processing*, Prentice Hall, Upper Saddle River, NJ, second edition, 176-182, (1999).

[2] SMITH, STEVEN W.: *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, 277-284, (1997).

[3] YOUNG, RICHARD A.: *The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles*, General Motors Research Laboratories, Vol. GMR-4920, 1-38, (1987).

[4] GONZALEZ, RAFAEL C. AND WOODS, RICHARD E.: *Digital Image Processing*, Pearson, New York, 4$^{th}$ edition, 167-176, (2018).

[5] DONOHO, DAVID L. AND JOHNSTONE, IAIN M.: *De-noising by soft-thresholding*, IEEE Transactions on Information Theory, Vol. 41, No, 3, 613-627, (1995).

[6] DONOHO, DAVID L. AND JOHNSTONE, IAIN M.: *Ideal spatial adaptation by wavelet shrinkage*, Biometrika, Vol. 81, No. 3, 425-455, (1994).

[7] JAZWINSKI, ANDREW H.: *Stochastic Processes and Filtering Theory*, Dover Publications, Dover reprint, New York, (2007).

[8] SÄRKKÄ, SIMO,: *Bayesian Filtering and Smoothing*, Cambridge University Press, Cambridge, (2013).

[9] JAZWINSKI, ANDREW H., *Stochastic Processes and Filtering Theory*, Dover Publications, Dover reprint, New York, (2007).

[10] ZIJAN ZHOU, YULI YANG, MENG MA AND BINGLI JIAC, *Distortion Reduction in Fractional Delay Filters*, IEEE Signal Processing Letters, PP (99):1-1, DOI:10.1109/lsp.2021.3064499, (2021).

[11] Y.SUN, Y. LIU, L. WANG, J.LI, J. WANG, A. ZHANG AND S. WANG, *Improved method for drone sound event detection system aiming at the impact of background noise and angle deviation*, Sens. Actuators, A 37, 115676, doi.org/10.1016/j.sna.2024.115676, (2024)