



# OPEN Flow prediction in sound-based uroflowmetry

Marcos Lazaro Alvarez<sup>1,3</sup>✉, Laura Arjona<sup>1,3</sup>, Mario Jojoa-Acosta<sup>2,3</sup> & Alfonso Bahillo<sup>2,3</sup>

Sound-based uroflowmetry (SU) offers a non-invasive alternative to traditional uroflowmetry (UF) for evaluating lower urinary tract dysfunctions, enabling home-based testing and reducing the need for clinic visits. This study compares SU and UF in estimating urine flow rate and voided volume in 50 male volunteers (aged 18–60), with UF results from a Minze uroflowmeter as the reference standard. Audio signals recorded during voiding were segmented and machine learning algorithms (gradient boosting, random forest, and support vector machine) estimated flow parameters from three devices: Ultramic384k, Mi A1 smartphone, and Oppo smartwatch. The mean absolute error for flow rate estimation were 2.6, 2.5 and 2.9 ml/s, with  $R^2$  values of 84%, 83%, and 79%, respectively. Analysis of the Ultramic384k's frequency range showed that the 0–8 kHz band contained 83% of significant components, suggesting higher sampling frequencies are unnecessary. A 1000 ms segment size was optimal for balancing computational efficiency and accuracy. Lin's concordance coefficients for urine flow and voided volume using the smartwatch (0–8 kHz, 1000 ms) were 0.9 and 0.85, respectively, demonstrating that SU is a reliable, cost-effective alternative to UF for estimating key uroflowmetry parameters, with added patient convenience.

**Keywords** Acoustic voiding signals, Flow prediction, Machine learning, Sound-based uroflowmetry

The rapid development of information and communication technologies is transforming healthcare systems, becoming remote and more proactive. This evolution brings benefits to both patients, by facilitating their access to medical services, and healthcare providers, who can obtain updated information and resources more quickly and efficiently. As a result, the quality of healthcare improves and associated costs decrease. Lower urinary tract symptoms (LUTS) are a problem affecting over 1900 million people worldwide<sup>1</sup>, leading to a decreased quality of life and significant healthcare resource expenditure<sup>2</sup>. The most widely used test to detect possible LUTS issues is UF<sup>3</sup>. UF is a non-invasive test based on estimating urine flow as a function of time, voided volume (VV), and the duration of the voiding process. This test is carried out in a clinic, where the patient must urinate into a uroflowmeter device. However, the accuracy of UF can be limited by various factors, such as an insufficient VV (which should be > 150 ml)<sup>4</sup> or the situational stress experienced by the patient due to the unnatural voiding environment. This process is often performed on demand, frequently with a low or very high bladder filling, rather than urinating when the patient feels physiologically ready<sup>5</sup>. Additionally, flow rates vary significantly throughout the day<sup>6</sup>, and a single test may not be representative of a patient's regular daily voiding patterns. This leads to non-reproducible flow measurements<sup>7</sup>, resulting in biased contributions from this test toward diagnosis<sup>8</sup>. Therefore, multiple tests are recommended, leading to numerous prolonged and costly visits for both the clinic and the patient<sup>1</sup>. To solve the challenges associated with UF, SU emerges to enable monitoring the void in a natural and comfortable environment for the patient, such as at home. This test seeks to estimate the flow patterns from the sound generated by urine hitting the water surface in the toilet bowl. Previous studies have shown a good correlation between the sound-derived flow parameters and those generated by UF<sup>8,9</sup>. To collect the audio signal for the SU tests, various hardware devices have been used in the literature. Smartphones are the most commonly used devices for performing SU tests due to their great versatility<sup>10–12</sup>. Other works have used professional microphones<sup>13</sup> and smartwatches<sup>14,15</sup> for the same purpose.

Recent studies have confirmed the potential of SU in estimating flow parameters, with strong correlations between SU and UF tests. In<sup>8</sup>, a moderate correlation with conventional uroflowmetry was reported, with a Pearson's correlation coefficient of 0.38, 0.57, and 0.68 for maximum flow rate (Qmax), average flow rate (Qave), and VV, respectively. In<sup>10</sup>, the authors reported strong correlations between Qmax, Qave, and VV of 0.88, 0.91, and 0.95 among men, and 0.78, 0.93, and 0.96 among women, respectively, for flow curve patterns obtained with their SU method and those obtained with a conventional UF device. However, the methods used to estimate flow parameters were not fully described and cannot be replicated. In<sup>16</sup>, the estimation of sound flow parameters was

<sup>1</sup>Faculty of Engineering, University of Deusto, 48007 Bilbao, Spain. <sup>2</sup>Department of Signal Theory and Communications, Universidad de Valladolid, 47002 Valladolid, Spain. <sup>3</sup>Marcos Lazaro Alvarez, Laura Arjona, Mario Jojoa-Acosta, and Alfonso Bahillo have contributed equally to this work. ✉email: alvarez.marcoslazaro@deusto.es

analysed. However, the methodology requires one to know the VV to obtain accurate estimates, which seems not feasible if we want to perform these tests at home. Finally, an analysis of the correlation of voiding characteristics was performed in<sup>12</sup> using deep learning on SU audio recordings. As an evaluation metric, the Lin concordance coefficient was used instead of the Pearson correlation coefficient applied in<sup>10</sup>. Although Lin's concordance coefficients were 0.77, 0.85, and 0.84 for Qmax, Qave, and VV respectively, lower than the corresponding Pearson coefficients obtained in<sup>10</sup>, the authors argue that Lin's concordance coefficient is more appropriate as an evaluation metric since it describes the degree of agreement between two measurements rather than the linear relationship given by Pearson.

The main objective of this research is to evaluate and estimate urine flow from SU tests using state-of-the-art machine learning (ML) regression algorithms. The comparative evaluation study considers different frequency bands in the range (0–96 kHz) and recording devices such as a Dodotronic Ultramic384, an Oppo smartwatch, and a Xiaomi Mi A1 smartphone. Furthermore, we used a professional Minze uroflowmeter<sup>17</sup> which allows us to obtain the actual flow for each time instant of the recorded audio signal. The Minze uroflowmeter, which serves as the ground truth, is factory calibrated and was purchased in November 2022. This device is a certified medical instrument that is compliant with the ISO 13485:2016 standard, ensuring highly accurate flow rate measurements. The data collected from this flowmeter was utilized to train the ML model.

The paper is organised as follows: “**Related work**” section briefly reviews the state of the art in audio feature extraction and flow prediction from SU audios using ML; “**Patients and methods**” section presents the materials and methods proposed in this research, describing the study design and the population, the characteristics of the datasets, and the procedures and theoretical foundations followed in the analysis of flow prediction in SU tests using different recording devices; “**Results**” section shows the results obtained from the proposed methodology; and finally, “**Conclusions**” section provides some concluding remarks.

## Related work

### Feature extraction in audio signals

There are multiple techniques in the literature to extract features from audio signals to be used in artificial intelligence models. Feature extraction techniques span several domains. The features of the temporal, frequency, cepstral, wavelet, and time-frequency domains have been explored for different types of audio signals, including speech, music, and environmental sounds<sup>18,19</sup>. The integration of modern ML algorithms with audio signal processing techniques has led to significant advances in audio classification tasks<sup>18</sup>. Studies have investigated the performance of different deep learning models using various audio features, such as the Mel Spectrogram and mel-frequency cepstral coefficients (MFCCs), both independently and in combination with ensembles<sup>20</sup>. The choice of features depends on both the dataset and the model, and feature combinations are generally restricted to those that perform well individually<sup>19</sup>.

In our research, we used the N linear-binned FFT of the audio signal to be analyzed, where N is an integer. The audio signal is processed using the Fast Fourier Transform (FFT) which is organised into a specific number of linearly distributed segments (or “bins”). Compared to other audio feature extraction techniques, such as cepstral analysis or MFCCs, N linear-binned FFT can be more intuitive and straightforward in its implementation. Furthermore, when combined with traditional ML regression methods, it offers a robust and efficient solution to extract and use audio features in our flow estimation analysis in SU tests with different recording devices.

### ML algorithms in SU

To estimate the voiding flow from SU tests, many algorithms have been implemented in the literature. Finding the optimal model and selecting the most suitable features from various audio segments can be a complex process. The proposed methods range from traditional signal processing techniques to more recent approaches utilizing deep learning. In<sup>11</sup>, ML was used with k-nearest neighbours and gradient boosting, and the model was trained using the MFCCs. In<sup>12</sup>, a deep neural network (DNN) with three hidden layers was used, taking spectral centroids, chroma vectors, and mel-frequency cepstral coefficients (MFCCs) as input features<sup>21,22</sup>. In<sup>23</sup>, an long short-term memory (LSTM) model was used for a time series prediction, with loudness representing the magnitude of the urinary sound and roughness representing the signal change pattern as inputs.

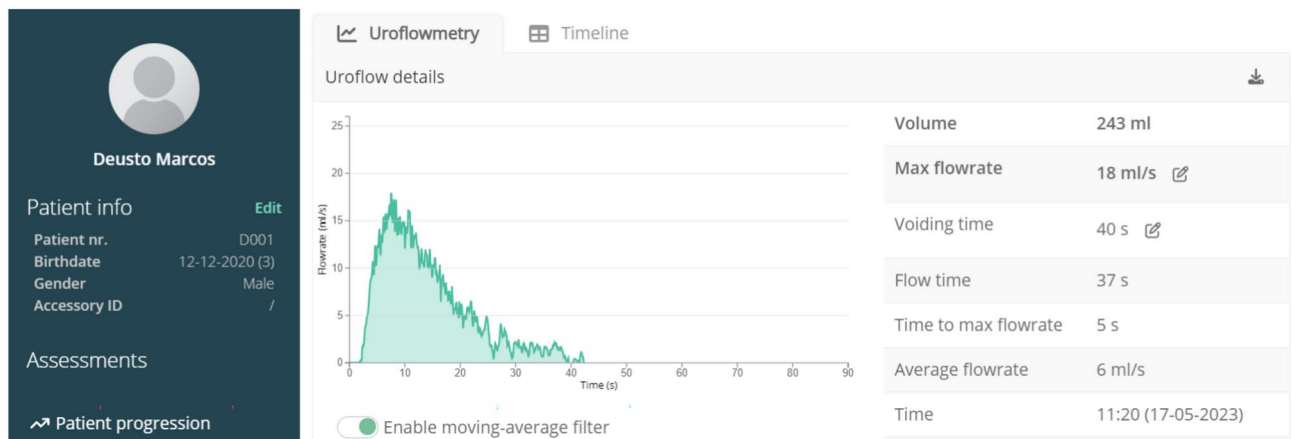
## Patients and methods

### Study design and population

We recruited 50 male volunteers without urological comorbidities, between the ages of 18 and 60, who agreed to participate in the study. The average height of the participants was found to be 175 cm, with a standard deviation of 4 cm. All participants provided their informed consent in writing for the use of conventional and acoustic uroflowmetry data. This study was approved by the Valladolid East Health Area Medicine Research Ethics Committee on 27 July 2023 with reference PI-GR-23-3275 (minutes number 16/2023). The Ethics Committee mentioned above complies with the GCP standards (CPMP/ICH/135/95).

### Procedures

The test consisted of urinating in a Minze uroflowmeter basin, while three different microphones recorded the sound. The Minze uroflowmeter basin, made of plastic, had been prefilled with 400 ml of water at the bottom of the container to ensure that the sound generated by the urination was against the water, simulating the conditions of a real toilet bowl. The Minze uroflowmeter has a resolution of 1 ml/s, with an accuracy of  $\pm 2.5$  ml/s for flow and  $\pm 30$  ml for VV<sup>17</sup>. The sampling rate is 10 Hz and records the flow curve, VV, Qmax, Qave, time to maximum flow and voiding time. Figure 1 shows an example of the data provided by Minze software during a test.



**Fig. 1.** Example of data recorded by the Minze uroflowmeter software during a test.

All participants were given instructions immediately before the test, asking them to aim at the toilet water as accurately as possible. After completing the test, they were asked if they had been able to meet this requirement, and all affirmative responses were subsequently validated by reviewing the audio recordings. Finally, recordings containing background noise, or those where there was uncertainty regarding time intervals in which urine may have hit the walls of the basin, were excluded from the analysis.

The entire voiding process on the Minze uroflowmeter was recorded using three sound recording devices:

- *Ultramic384 (UM)*: a high-quality microphone (FG23629 microphone sensor from Knowles), used in studies to record void events SU<sup>14</sup> with a sampling rate (SR) of 384 kHz, allowing the study of a wide frequency spectrum. For our tests, we used a SR of 192 kHz, because tests showed that there was no information above the 96 kHz band. This device is not intuitive to use as it requires additional hardware components to operate. Besides, it is not as versatile compared to smartwatches and smartphones (the device has no other use beyond audio recording).
- *Mi A1 smartphone (Phone)*: it integrates a medium-quality microphone (SPU0410LR5H-QB microphone sensor from Knowles), with a SR of 44.1 kHz, and it has also been used in similar studies. This device is intuitive to use, although not as versatile during urination as a smartwatch<sup>24</sup>.
- *Oppe smartwatch (Watch)*: it integrates a medium-quality microphone (chipset details not publicly disclosed), validated for use in SU applications<sup>14</sup>, with a SR of 44.1 kHz. It is the most intuitive and versatile recording device for the urination process. During the test, the Watch and the UM were placed opposite the user at a height of 80 cm above the basin water level, that is the average height of a toilet cistern. The test participants must wear the smartwatch on the wrist, pointing its microphone to the basin during the recording. Figure 2 shows the laboratory environment set up in the bathroom where the tests were performed. The bathroom had dimensions of 404 cm (length) × 175 cm (width) × 271 cm (height), with ceramic tiles covering the floor and walls, and a plasterboard ceiling.

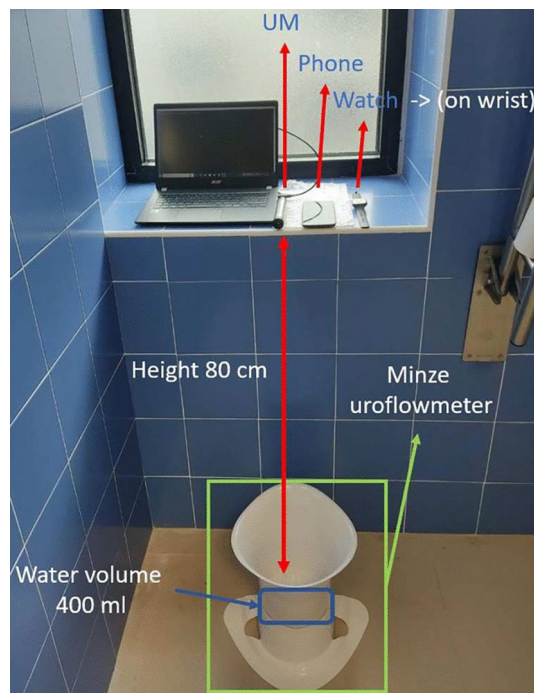
### Dataset description

For each performed test, we obtained three audio files in WAV format, corresponding to each of the recording devices. For each test, the corresponding uroflowmetry curve provided by Minze software, containing the flow information for each time unit (100 ms resolution), was matched to these audio files.

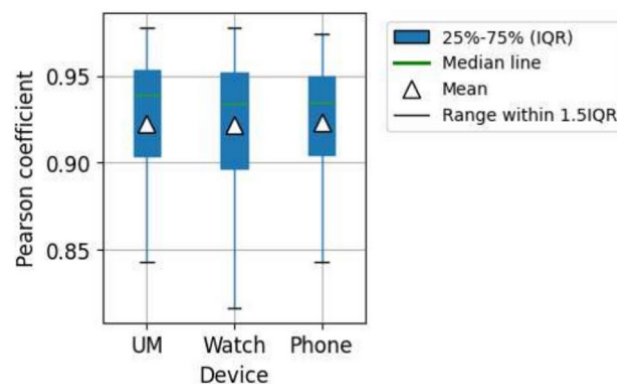
After excluding the test that presented the issues mentioned in “[Patients and methods](#)” section, we obtained a total of 47 valid tests. For each test, we obtained the corresponding flow values with a frequency of 10 Hz for further analysis.

### Correlation between SU signal envelopes and Minze measurements

To study the sound signals’ waveform and compare it with the flow graphs provided by Minze software, we obtained the sound envelope from each one of the recorded audios following the methodology proposed by<sup>14</sup>. We then evaluated the linear relationship between the flow envelope given by Minze software and the sound signal envelope of the three recording devices. We used the Pearson correlation coefficient to analyze the 47 audio signals recorded by each of the devices. We selected Pearson’s correlation coefficient because we are interested in the linear relationship rather than the level of similarity. Figure 3 shows the result of evaluating the envelope correlation using a box plot. It can be seen that the interquartile range (IQR) for the three devices is very close to the median and mean and is it relatively narrow. This suggests that the correlations are consistently high and that there is little variability in the correlation between the waveform provided by Minze and the envelope of the sound signal for each device, supporting the idea of estimating flow from sound in SU tests. Figure 4 shows a random selection of eight tests in which a very high correlation is demonstrated, further supporting the idea of estimating flow from sound using SU tests.



**Fig. 2.** Laboratory data collection scenario showing the Minze uroflowmeter with a water volume of 400 ml in the basin and the three recording devices: UM, Phone, and Watch, along with their respective heights.



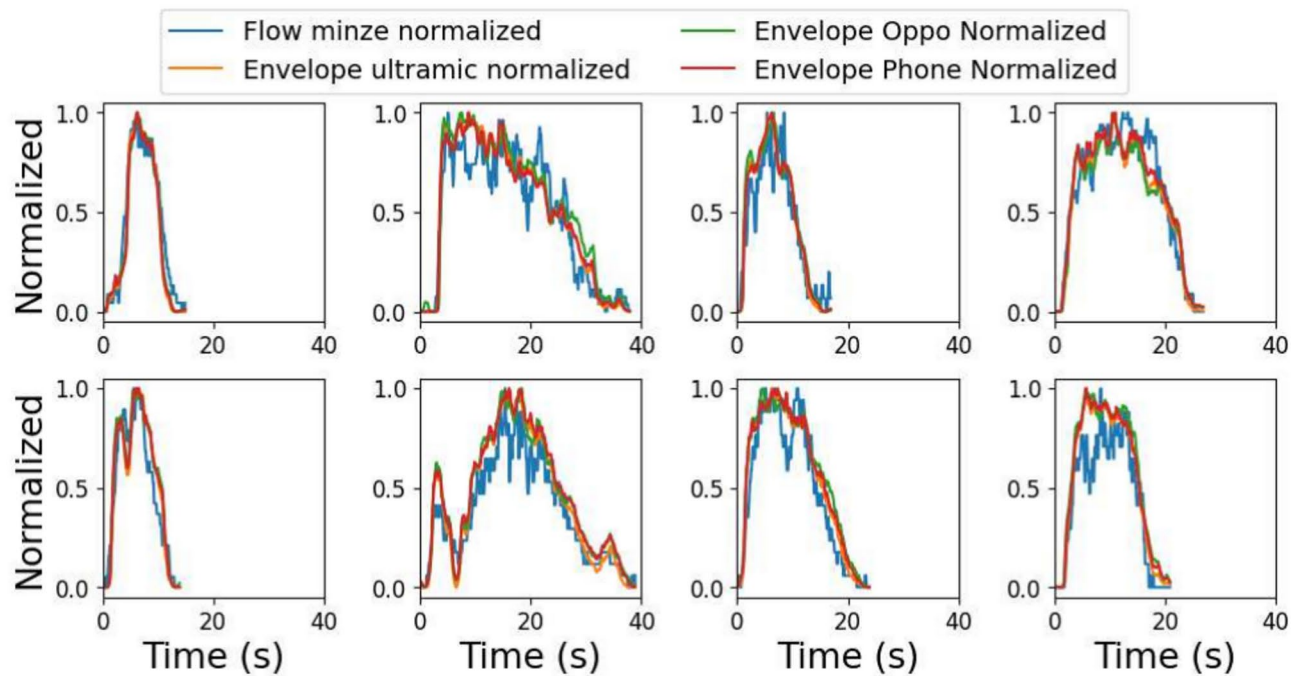
**Fig. 3.** Results of evaluating Pearson's correlation for the three devices to compare the envelop extraction.

### General research diagram

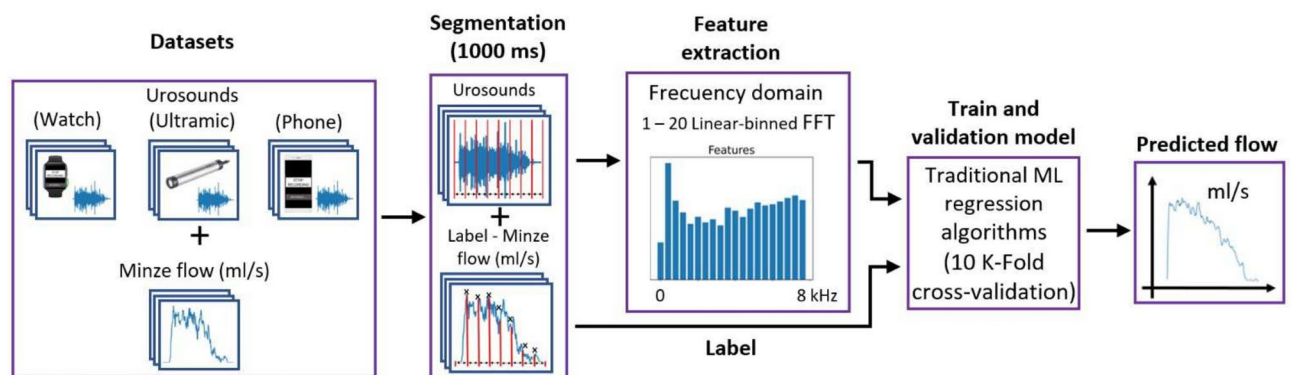
Figure 5 shows a graphical diagram of the proposed methodology for analyzing SU audios to predict the voiding flow rate. Our input data are the SU audio signals in WAV format recorded with the three devices (see “[Patients and methods](#)” section). Each audio is associated with its corresponding flow data point provided by the Minze software. To analyse the audio signals, the first step was to synchronize all the signals with the data given by Minze. The audio signals were then split into 100-ms segments and each segment was labelled with the flow value provided by the Minze software. We selected a segment size of 100 ms because the Minze software provides flow data points sampled at 10 Hz. As a result of the audio segmentation, we obtained 13,060 100-millisecond-long segments of SU labelled with their corresponding flow value.

To estimate the flow from sound, we used traditional ML regression algorithms. The input features for training the algorithms were frequency domain features. We utilized the N linear-binned fast fourier transform (FFT), where the frequency range is divided into N equally spaced intervals. For each interval, we sum the absolute values of the amplitudes of the components present in each interval, obtaining a vector with N values that characterize each audio segment. Linear-binned FFT is a widely used technique to extract the most relevant features of the frequency spectrum of an audio signal and has demonstrated effectiveness for audio analysis. For our experiments, we selected  $N = 20$ , a value chosen based on experimentation with various values of N.





**Fig. 4.** Comparison of Minze flow data with the sound envelope of the signal given by UM, Watch, and Phone, for a selection of eight randomly selected signals.



**Fig. 5.** Diagram showing the pipeline of the proposed methodology, consisting of 4 main steps: data extraction, audios segmentation, feature extraction, and finally model training and validation.

The performance of the regression models was evaluated using tenfold cross-validation. In this process, the dataset is divided into 10 subsets; in each iteration, one subset is used as the test set, while the other nine are used for training. The model's performance is assessed on the test set for each fold, and the final results are reported as the average error across all tenfolds. This ensures that all data is used for both training and testing, but never simultaneously, providing an unbiased estimate of the model's predictive capability.

To predict the flow rate of the labelled audio segments, we performed an analysis using three ML regression algorithms, with the aim to evaluate which one performs best in the flow prediction task. Below are the details of the algorithms used:

- *Random forest regressor (RF)*: is widely used in various fields such as healthcare, where the prediction of continuous values is essential. It is a versatile and powerful algorithm known for its effectiveness in a variety of regression tasks.
- *Support vector regressor (SVR)*: is suitable for tasks where capturing non-linear relationships is essential, and it can be a powerful tool in various regression applications, including time-series analysis.
- *Gradient boosting regressor (GBR)*: is widely used in practice for tasks such as time-series forecasting and various regression applications where accurate predictions are essential.

Figure 6 shows the evaluation results of different regression algorithms in terms of their MAE for each one of the recording devices, with a segment duration of 100 ms. The MAE is defined as the average of the absolute differences between the predicted and true flow values, computed across all labeled audio segments, regardless of the flow magnitude. To train our ML algorithms, we used tenfold cross-validation due to its multiple advantages in terms of model performance evaluation. We used a value of  $N = 13,060$ , which corresponds to the number of labelled 100-ms audio segments. It can be observed that, in general, the three algorithms show similar results, with RF performing slightly better.

For our remaining analysis, we used RF because it presents a compromise between accuracy and implementation complexity.

### Feature selection

The first step to train our regression model is to select the best procedure for characterizing each audio sample in the dataset. First, we performed an analysis across the entire frequency band recorded by the professional UM microphone (0–96 kHz) to identify which frequency components have the most influence on the flow prediction. For this purpose, we extracted 1000 linear-binned FFT samples for each 100-ms audio segment. As a result, we obtained a vector with 1000 values that characterize each audio segment. We trained a supervised RF algorithm using the 1000 linear-binned FFT samples for each labelled audio segment with the corresponding flow label provided by Minze. Subsequently, from the trained RF model, we obtained a Gini impurity-based metric<sup>25</sup> to measure the quality of our split criteria. This metric allows us to quantify the weighted impurity of each feature in the tree, indicating its relative importance in the model. Furthermore, the Gini impurity metric provides us with an effective way to interpret the model and understand which frequencies are most influential in the predictions.

Figure 7 shows that the frequency components that contribute the most information to the prediction of flow from sound are in the lower band of the signal spectrum. The upper part of the spectrum above 8 kHz is not relevant for our application. For the remaining analyses, we used the (0–8 kHz) band, which contains 83% of the frequency components that contribute the most information to the model, and is the band in which the majority of commercial devices record.

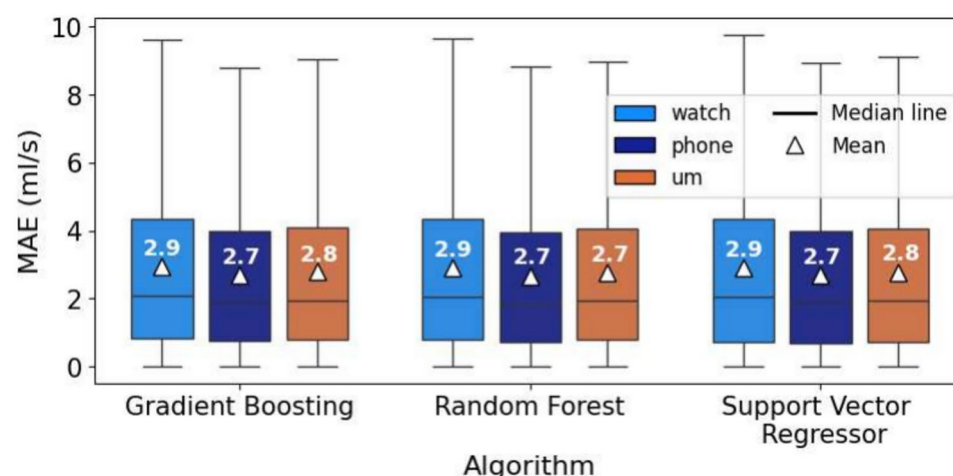
### Analysis of different audio segment sizes

Once we analyzed in “Patients and methods” section the frequency bands that have a higher influence on the model prediction results, we then performed an analysis to study the influence of the audio segment size taking the 0–8 kHz frequency band. The Minze uroflowmeter used has a SR of 10 Hz, so the maximum resolution we can achieve for the audio signals is a segment duration of 100 ms.

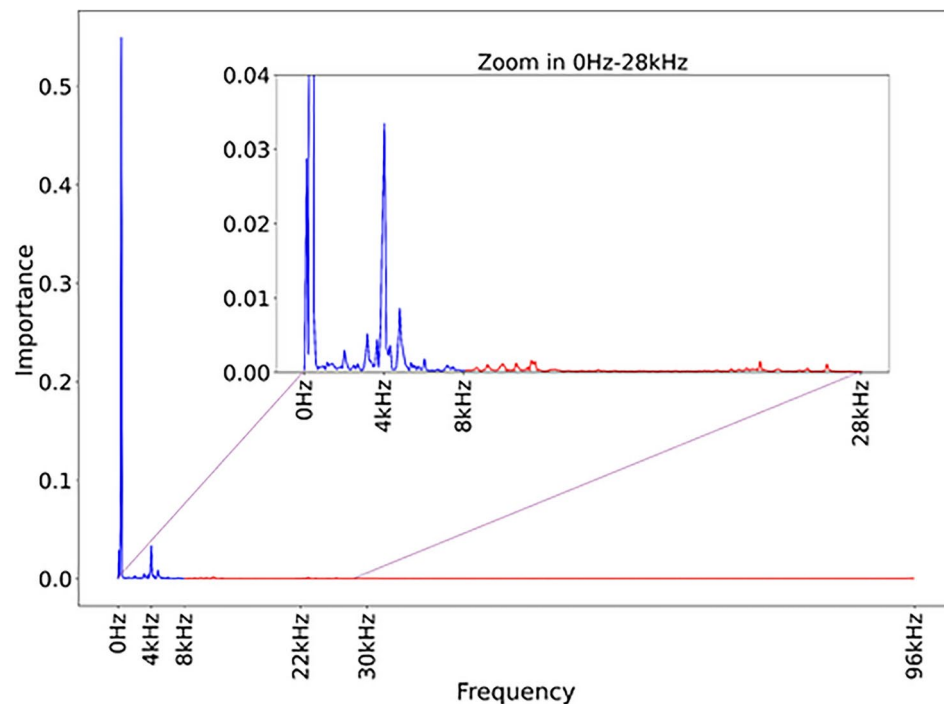
For each recording device, we analyzed eight different audio segment sizes: 100, 200, 500, 800, 1000, 1100, 1200, and 1500 ms. For segments longer than 100 ms, we took the maximum flow value provided by Minze for the corresponding timestamp as the flow label. We selected this value because, according to urologists, the most valuable information in a uroflowmetry test is the maximum flow reached by the patient at any given moment. For this study, we used a RF model taking as input features the 20 linear-binned FFT samples for each audio segment.

Figure 8 shows the evaluation results in terms of MAE. Across all devices, the segment size that yielded the lowest MAE, indicating the best performance, was 1000 ms. The audio recordings from the Phone and UM showed similar and lower errors compared to those from the Watch device.

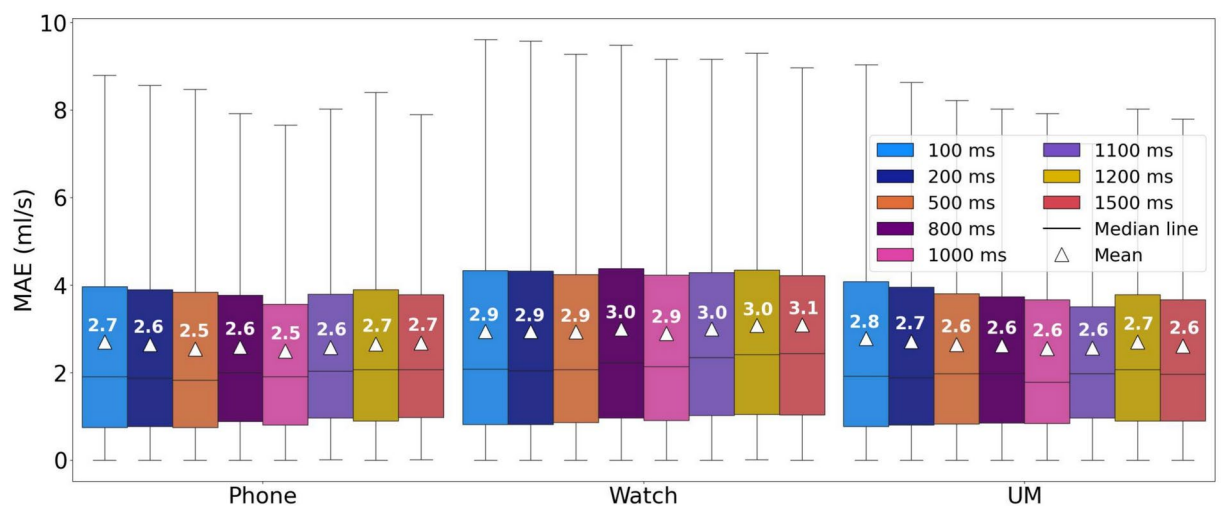
Despite the Watch exhibiting slightly higher errors than the other two devices (on average 0.37 ml/s higher than the phone and 0.30 ml/s higher than the UM), it was selected for further analysis. This decision was driven



**Fig. 6.** Evaluation results of the three regression algorithms for each recording device, in terms of the MAE, measured in ml/s. The MAE is the mean of the absolute differences between the predicted and actual flow values, calculated across all labeled audio segments.



**Fig. 7.** Predictive power (importance) of each frequency component in the flow prediction task from SU. The frequency band selected in our algorithms is shown in blue, showing the highest values of importance. Importance is calculated using the Gini impurity metric with an RF model.



**Fig. 8.** Analysis of the MAE for the RF prediction model, comparing different audio segment sizes (ms) and the three different recording devices. The MAE is the mean of the absolute differences between the predicted and actual flow values, calculated across all labeled audio segments.

by Watch's versatility and ease of use within the SU context, which offsets the slight increase in error. Figure 8 showed that the error increase of the Watch compared to the other devices is not significant.

### Frequency analysis

Once we have determined the audio segments size that yields the best prediction results, we then analyse the frequency range within the (0–96 kHz) band that also obtains the best flow estimation results. This study is performed using the audios recorded with the UM with a audio segment size of 1000 ms. We train a RF model utilising the 20 linear-binned FFT as input features.

To carry out this analysis, we considered the following points:

- We eliminated the low-frequency noise bands (0–250 Hz) and (0–1200 Hz). These frequency ranges have been adopted from related works<sup>14,16</sup>, respectively.
- To study the flow prediction results in environments where the user privacy preservation is required, we considered the removal of three frequency bands: (0–4 kHz) and (0–8 kHz) that correspond to the conversational band, and (0–16 kHz) that corresponds to the human audible range.
- An analysis was performed within the (0–8 kHz) band because it contained the most information, as shown in the analysis performed in “Patients and methods” section. In addition, this frequency range is used by the vast majority of commercial devices. This selection represents a compromise between the performance of the model and the cost and availability of the microphone.

Figure 9 shows the results of evaluating the RF model for 18 different frequency bands, from 0 to 96 kHz. The best results were obtained for the (0–8 kHz) and (0–5 kHz) bands, demonstrating that most of the prediction power for the flow prediction task is found in the lower end of the spectrum. It can be observed that when the speech band (0–4 kHz) is removed, the error increases. However, it could be an alternative, despite being less accurate, for applications where preserving privacy is crucial.

## Results

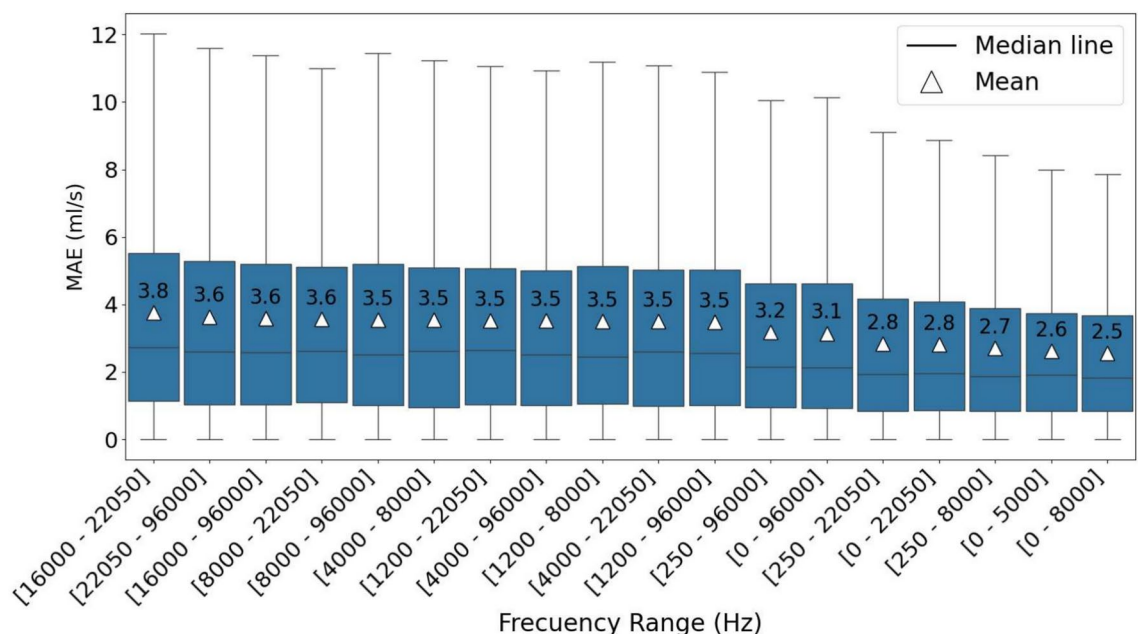
### Flow estimation with selected parameters

The voiding flow is the variable that provides the most valuable information for urologists from a flowmetry test. To predict the flow using audio signals, in this section we train a new RF model using the parameter settings selected in the analyses performed in the previous sections.

We trained a RF model using audio signals recorded from the Watch, segmented into 1000-ms segments, within the band of 0–8 kHz. As features, we used the 20 linear-binned FFT. To mitigate the effect of overfitting, we used k-fold cross-validation, as it provides a robust and reliable estimate of a model's performance on unseen data. For our algorithms, we selected  $k = 10$  to achieve a reasonable balance between bias and variance in the performance estimate. Figure 10 presents the results obtained with a MAE of 2.86 ml/s. We can consider that the error obtained by our algorithm is significantly low, since our reference device, the Minze uroflowmeter, according to the manual introduces a base error of  $\pm 2.5$  ml/s. To assess the correlation between the predicted and actual values, we applied Lin's concordance correlation coefficient for each 1000-ms segment, yielding a value of 0.9. This high coefficient suggests a strong positive relationship, indicating that, as voiding flow increases, the model's predictions also increase in a manner consistent with the field measurements.

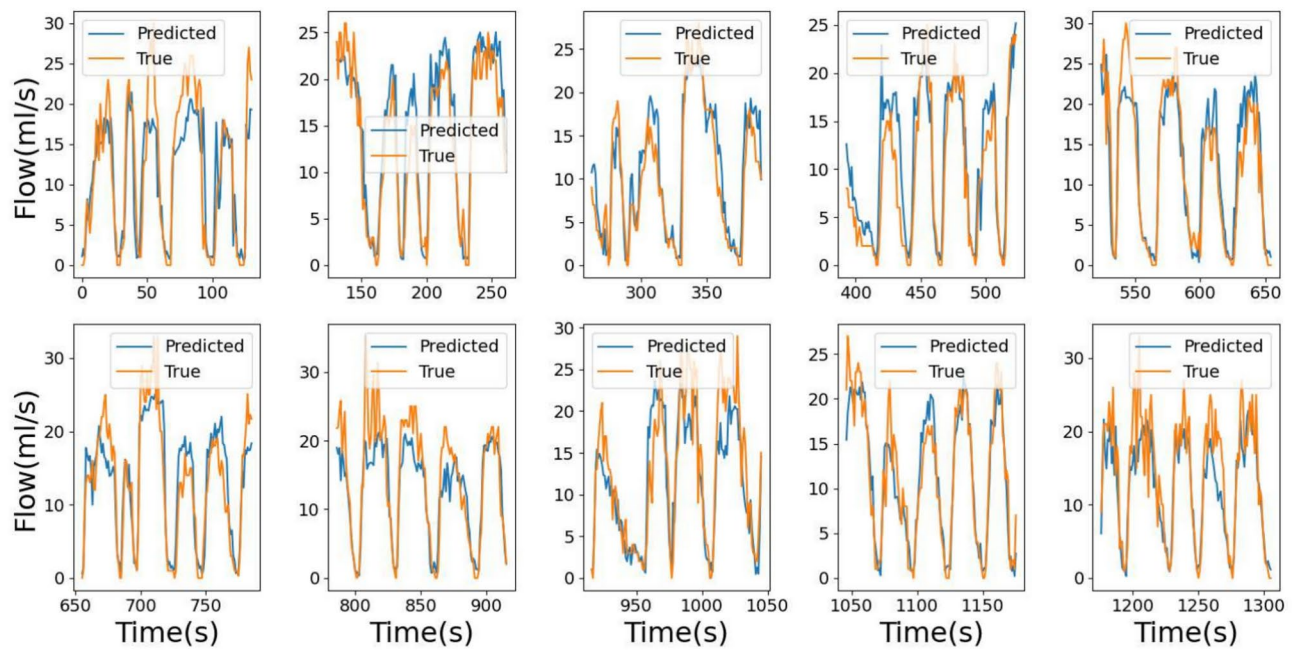
### Voiding volume estimation with selected parameters

Another important parameter for urologists to evaluate the urinary tract is the VV. In our dataset, we also have the VV associated with each of the audio recordings, obtained from the Minze uroflowmeter. To estimate the VV for each test, we sum the estimated flows corresponding to each audio signal from the SU tests, as obtained in “Results” section. We evaluated Lin's concordance correlation coefficient between the predicted and actual

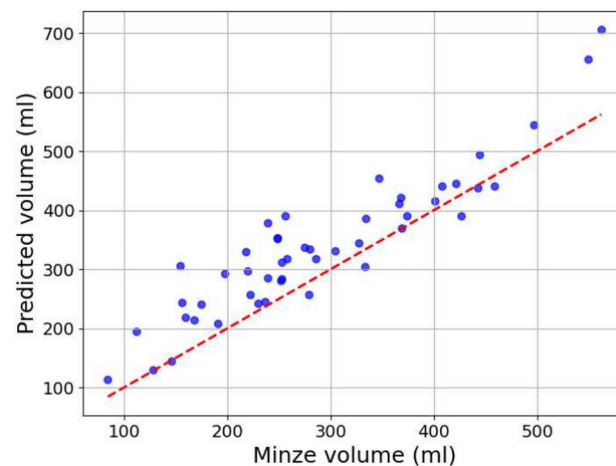


**Fig. 9.** Analysis of the RF model MAE value for different frequency bands, using the UM microphone. The MAE is the mean of the absolute differences between the predicted and actual flow values, calculated across all labeled audio segments.





**Fig. 10.** Comparison of the UF (orange) and SF (blue) flow curves. To obtain the SF curves, we selected the Watch audios, and the model selected was a RF with 20 linear-binned FFT as input features, taking segment sizes of 1000 ms.



**Fig. 11.** Comparison between the VV given by Minze flowmeter and the predicted volume, calculated as the sum of the estimated flows corresponding to each audio signal from the SU tests.

volume values and obtained a value of 0.85. Figure 11 shows the result of the comparison between the estimated and actual volumes provided by Minze uroflowmeter.

## Conclusions

The comparative results to conduct SU tests using three different recording devices showed similar performance for the Watch and UM, and slightly lower performance for the Watch, for the task of estimating the flow rate. However, the choice of the Watch is based on its ease of use and versatility for conducting UF tests at home. Also, the estimation error difference across the three devices is not significant. Moreover, smartwatches, being wearable objects unlike the other microphones, are more suitable for these types of tests because:

- They are very comfortable for the user as they do not interfere with the act of voiding, can be activated automatically without requiring any action from the patient (especially useful for individuals with low digital skills, such as the elderly and children).

- Since they have a fixed position on the user's body, they allow for sound recording from a consistent distance from where the sound is produced.
- They can be used continuously, making it possible to maintain a voiding diary that measures multiple flows at different times of the day and night, which could be considerably more useful and objective for determining any pathophysiology.

These characteristics position the Watch as a more versatile and user-friendly device for extended home monitoring compared to handheld alternatives such as the Watch, while providing significantly high accuracy to estimate the flow rate.

Contrarily, the UM microphone is not a wearable device, requires additional data collection equipment, is not user-friendly, is bulky, and is not widely distributed, which could limit its use in SU. Its use is only justified if one seeks to obtain information in frequencies above 22 kHz (ultrasound applications), which is particularly beneficial for applications where it is necessary to preserve privacy by eliminating the audible human band.

## Limitations

- *Environmental variability* The generalization of the algorithm across different environments has not been evaluated. Variations in environment, as well as differences in toilet material, size, and shape, could potentially affect the acoustics of urination event. To address these variations, additional data collection in diverse settings is necessary to refine and calibrate the algorithm for different conditions.
- *Gender differences* Our study primarily enrolled male participants due to the differences in voiding habits between men and women. Factors such as standing versus sitting position influence how the urine stream impacts the toilet bowl, thereby altering the acoustics of the urination process. As part of future work, we intend to collect a new dataset that includes female volunteers, aiming to enhance the model's applicability across diverse patient demographics.

## Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 21 October 2024; Accepted: 30 December 2024

Published online: 03 January 2025

## References

1. Alothmany, N. et al. Critical review of uroflowmetry methods. *J. Med. Biol. Eng.* **38**, 685–696 (2018).
2. Arjona, M. F. & Sanz, I. P. Hiperplasia benigna de próstata: Una afección de elevada prevalencia en el paciente de edad avanzada. *Rev. Española Geriatr. Gerontol.* **43**, 44–51 (2008).
3. Sorel, M. R., Reitsma, H. J., Rosier, P. F., Bosch, R. J. & de Kort, L. M. Uroflowmetry in healthy women: A systematic review. *Neurourol. Urodyn.* **36**, 953–959 (2017).
4. Rodrigues Pessoa, R. & Kim, F. *Urodynamics and Voiding Dysfunction* 7th edn, 452–454 (Abernathy's Surgical Secrets, 2018).
5. Kuoch, K. L., Meyer, D., Austin, D. W. & Knowles, S. R. Classification and differentiation of bladder and bowel related anxieties: A socio-cognitive exploration. *Curr. Psychol.* **40**, 4004–4011 (2021).
6. Golomb, J., Lindner, A., Siegel, Y. & Korczak, D. Variability and circadian changes in home uroflowmetry in patients with benign prostatic hyperplasia compared to normal controls. *J. Urol.* **147**, 1044–1047 (1992).
7. Sonke, G. S. et al. A method for estimating within-patient variability in maximal urinary flow rate adjusted for voided volume. *Urology* **59**, 368–372 (2002).
8. Krhut, J. et al. Comparison between uroflowmetry and sonouroflowmetry in recording of urinary flow in healthy men. *Int. J. Urol.* **22**, 761–765 (2015).
9. Gärtner, M. et al. Evaluation of voiding parameters in healthy women using sound analysis. *LUTS Lower Urin. Tract Symptoms* **10**, 12–16 (2018).
10. Lee, Y. J., Kim, M. M., Song, S. H. & Lee, S. A novel mobile acoustic uroflowmetry: Comparison with contemporary uroflowmetry. *Int. Neurourol. J.* **25**, 150 (2021).
11. Aslim, E. J. et al. Pilot study for the comparison of machine-learning augmented audio-uroflowmetry with standard uroflowmetry in healthy men. *BMJ Innov.* **6**, bmjinnov-2019-000382 (2020).
12. Lee, H. J. et al. Development and validation of a deep learning system for sound-based prediction of urinary flow. *Eur. Urol. Focus* **9**, 209–215 (2023).
13. Alvarez, M. L., Arjona, L., Iglesias Martínez, M. E. & Bahillo, A. Automatic classification of the physical surface in sound uroflowmetry using machine learning methods. *EURASIP J. Audio Speech Music Process.* **2024**, 12 (2024).
14. Arjona, L., Díez, L. E., Bahillo, A. & Arruza-Echevarria, A. Urosound: A smartwatch-based platform to perform non-intrusive sound-based uroflowmetry. *IEEE J. Biomed. Health Inform.* **27**, 2166–2177 (2022).
15. Narayanswamy, G., Arjona, L., Díez, L. E., Bahillo, A. & Patel, S. Automatic classification of audio uroflowmetry with a smartwatch. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 4325–4329 (IEEE, 2022).
16. El Helou, E. et al. Mobile sonouroflowmetry using voiding sound and volume. *Sci. Rep.* **11**, 11250 (2021).
17. Homepage - Minze Health, accessed on 28 October 2024; <https://minzehealth.com/> (2024).
18. Sharma, G., Umapathy, K. & Krishnan, S. Trends in audio signal feature extraction methods. *Appl. Acoust.* **158**, 107020 (2020).
19. Alias, F., Socoró, J. C. & Sevillano, X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Appl. Sci.* **6**, 143 (2016).
20. Turab, M., Kumar, T., Bendechache, M. & Saber, T. *Investigating multi-feature selection and ensembling for audio classification*. arXiv preprint [arXiv:2206.07511](https://arxiv.org/abs/2206.07511) (2022).
21. Piczak, K. J. Environmental sound classification with convolutional neural networks. In *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)* 1–6 (IEEE, 2015).
22. Balamurali, B. et al. Deep neural network-based respiratory pathology classification using cough sounds. *Sensors* **21**, 5555 (2021).
23. Jin, J., et al. Prediction of urinal flowrate by using urinary acoustic signals based on LSTM neural network. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* 5807–5810, vol. 261 (Institute of Noise Control Engineering, 2020).

24. Castillo, Y., Cámara, M. A., Blanco-Almazán, D. & Jané, R. Characterization of microphones for snoring and breathing events analysis in mhealth. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 1547–1550 (IEEE, 2017).
25. Iravantchi, Y., Ahuja, K., Goel, M., Harrison, C. & Sample, A. Privacymic: Utilizing inaudible frequencies for privacy preserving daily activity recognition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* 1–13 (2021).

## Acknowledgements

This research was supported by the Spanish Ministry of Science and Innovation under SWALU project (ref. CPP2022-010045) and 'Ayuda para contratos predoctorales 2020 (ref. PRE2020-095612)' funded by MICIU/AEI /10.13039/501100011033 and co-financed by FSE invierte en tu futuro. Additionally, partial support was provided by the Ministry under the Aginplace project (ref. PID2023-146254OB-C41 and ref. PID2023-146254OA-C44).

## Author contributions

M.L.A., A.B., and L.A. conceived the study. M.L.A. and L.A. designed the experiments, while M.J.-A. developed the theoretical framework. M.L.A. performed the computations and, together with A.B., contributed to sample preparation and data collection. All authors participated in the analysis and interpretation of the results, and contributed to the writing and revision of the manuscript. M.L.A. drafted the manuscript, incorporating feedback from all authors.

## Declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.L.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, corrected publication 2025