



Data Article

IMU50: a high-frequency multi-modal wrist-worn IMU dataset of 50 subjects over 5 continuous days for supervised and self-supervised learning in free-living conditions



Irene La Fratta^a, Lorenzo Calisti^{b,c}, Leonardo Bigelli^{b,c},
Piero Chiacchiarretta^{d,e}, Alessandra Mascitelli^{d,e,f}, Damiano D'Ardes^g,
Piero Di Carlo^{d,e}, Francesco Lucertini^h, Antonio Ferretti^{a,i,l,*},
Emanuele Lattanzi^{b,c,*}, Raffaella Franciotti^{a,i,l}

^aDepartment of Neuroscience, Imaging and Clinical Sciences, University "G. d'Annunzio" of Chieti-Pescara, Italy

^bDepartment of Pure and Applied Sciences (DiSPeA), University of Urbino, Italy

^cArtificial Intelligence Multidisciplinary Research Center (AIMRC), University of Urbino, Italy

^dDepartment of Advanced Technologies in Medicine & Dentistry, University "G. d'Annunzio" of Chieti-Pescara, Italy

^eCenter for Advanced Studies and Technology (CAST), University "G. d'Annunzio" of Chieti-Pescara, Italy

^fInstitute of Atmospheric Sciences and Climate (ISAC), National Research Council (CNR), Via del Fosso del Cavaliere, 100, Rome, 00133, Italy

^gDepartment of Medicine and Aging Science, University "G. d'Annunzio" of Chieti-Pescara, Italy

^hDepartment of Biomolecular Sciences (DiSB), University of Urbino, Italy

ⁱInstitute for Advanced Biomedical Technologies (ITAB), University "G. d'Annunzio" of Chieti-Pescara, Italy

^lUdA-TechLab, Research Center, University "G. d'Annunzio" of Chieti-Pescara, Italy

ARTICLE INFO

Article history:

Received 27 July 2026

Revised 5 August 2026

Accepted 22 August 2026

Available online 27 August 2026

ABSTRACT

We present a multiday, high-resolution inertial measurement unit (IMU) dataset collected from 50 healthy volunteers using the Actigraph Leap wrist-worn device. Each participant was monitored continuously for five consecutive days and nights, yielding an average of approximately 136 h of raw sensor recordings per subject. The dataset includes three-axis accelerometer signals, three-axis gyroscope signals sampled at 128 Hz, double-channel photoplethysmogram (PPG) sampled at 25 Hz, and wrist skin temperature acquired every minute.

* Corresponding authors.

E-mail addresses: antonio.ferretti@unich.it (A. Ferretti), emanuele.lattanzi@uniurb.it (E. Lattanzi).

Dataset link: [IMU50: a high-frequency multi-modal wrist-worn IMU dataset of 50 subjects over 5 continuous days for supervised and self-supervised learning in free-living conditions \(Original data\)](#)

Keywords:
Inertial measurement unit
Wearable sensor
Daily-living activity monitoring
Supervised learning
Self-supervised learning
Energy expenditure
Morphological metadata

Hourly estimates of metabolic equivalents of task (METs) and caloric expenditure are also provided. These estimates are obtained by the Actigraph using Freedson algorithms computed over an additional 32 Hz embedded accelerometer. In addition to sensor data, the repository contains participant-level metadata encompassing age, sex, body weight, height, body mass index (BMI), hip and waist circumference, educational level, and lifestyle. All recordings were conducted under free-living conditions without restrictions on daily activities, capturing a realistic spectrum of behaviours including sleep, sedentary periods, and light daily activities. Data are stored in open, well-documented formats with accompanying code examples to facilitate loading and preprocessing. The dataset is intended to support research in two complementary directions: (1) self-supervised representation learning from raw IMU signals, exploiting the large volume of unlabelled continuous recordings; and (2) supervised regression and classification tasks targeting metabolic outcomes, morphological parameters, skin temperature, and heart rate from PPG data. No activity labels were collected during recording, making this resource particularly suitable for unsupervised and weakly supervised learning paradigms. The dataset fills a gap in the public availability of long-duration, multi-day wrist IMU recordings linked to metabolic and clinical metadata and is expected to serve as a benchmark resource for the wearable computing and digital health communities.

© 2026 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Specifications Table

Subject	Please select Subject from the dropdown list. Computer Sciences
Specific subject area	Wearable sensors, daily-living activity monitoring, digital health, machine learning, and Human Activity Recognition
Type of data	Time-series signal files containing Raw IMU signals at 128 Hz (CSV per subject); dual-channel PPG signals at 25 Hz (CSV per subject); skin temperature measurements every minute (CSV per subject); hourly MET and caloric expenditure estimates (CSV per subject); metadata table (CSV)
Data collection	The study was conducted in accordance with the Declaration of Helsinki and approved by the Biomedical Research Ethics Committee of “C.Et.R.A.—Comitato Etico Regione Abruzzo” (protocol code 1782 of 8 May 2023). Participants aged 40 to 75 years were recruited through public advertisements and had no history of motor impairment, neurological, cardiovascular, and respiratory diseases, and provided written informed consent. Participants were asked to wear the ActiGraph LEAP wrist-worn device on the non-dominant wrist for five consecutive days and nights under free-living conditions. Sensor data were collected together with demographic, anthropometric, and lifestyle information.
Data source location	The collected data are part of the VITALITY project (“Innovation Ecosystem: innovation, digitalisation and sustainability for the widespread economy in central Italy”), funded by NextGenerationEU, within the activities of the Spoke of the G. d’Annunzio University “Telemedicine One-Health and Environment”. The present data were collected at the Department of Neuroscience, Imaging and Clinical Sciences (DNISC), G. d’Annunzio University of Chieti-Pescara, Chieti, Italy. The Dataset is stored at the Zenodo repository.

(continued on next page)

Data accessibility	Repository name: Zenodo Data identification number: 10.5281/zenodo.21,468,410 Direct URL to data: https://zenodo.org/records/21468410 Instructions for accessing these data: The dataset is publicly accessible without registration. Files can be downloaded directly from the Zenodo record page.
Related research article	None

1. Value of the Data

- This dataset provides one of the largest publicly available collections of continuous, multiday wrist-worn IMU recordings linked to metabolic outcomes, demographic, anthropometric, and lifestyle metadata. The combination of 128 Hz raw signals, hourly metabolic estimates, and participant data in a single cohort is rarely available in open repositories.
- The dataset enables self-supervised and unsupervised learning from raw IMU signals without the need for activity labels. Researchers can develop and benchmark masked autoencoding, contrastive learning, and temporal prediction tasks on a realistic, valid corpus of free-living human movement data.
- The hourly MET and caloric expenditure provide supervised regression targets for energy expenditure modelling from raw IMU signals, supporting research into metabolic inference applicable to wearable health monitoring.
- The demographic (age, sex, education), anthropometric (body weight, height, circumference of the hips and waist, BMI), lifestyle (smoke and alcohol), and heart rate from PPG allow researchers to explore supervised prediction of individual characteristics from sensor raw signals and to investigate confounding effects of included variables on IMU-derived features.
- Continuous night-time recordings of IMU and PPG enable sleep and circadian rhythm analysis without polysomnography, supporting actigraphy-based sleep research and the development of wrist-movement proxies for sleep quality metrics in free-living populations.
- The dataset can serve as a transfer learning source domain for downstream tasks in digital health, rehabilitation, and sports science, where labelled data are scarce. Pre-trained representations derived from this corpus may generalise to activity recognition, fall detection, and gait analysis benchmarks.

2. Background

The growing availability of low-cost inertial measurement units (IMUs) embedded in consumer-grade wearable devices has stimulated interest in the automated inference of physical activity, energy expenditure, and health-related outcomes from continuous movement data [3,4]. However, progress in this field remains constrained by the scarcity of publicly available datasets that combine individual-level metadata with high-resolution, multimodal raw IMU signals with concurrent physiological references collected under free-living conditions over multiple days [5,6]. To contextualize this gap, Table 1 provides a comprehensive comparison between the dataset presented in this study and the most prominent publicly available benchmarks.

Among these, CAPTURE-24 [7] provides wrist accelerometer recordings from 151 participants collected over approximately 24 h per subject in unconstrained, real-world settings, resulting in one of the largest annotated free-living datasets currently available. Resources such as the UK Biobank [8] and the China Kadoorie Biobank (CKB) [9] offer unlabelled data at a much larger population scale. Both repositories rely on wrist-worn accelerometers worn continuously for seven days and involve tens of thousands of participants recruited from the general population. While neither provides activity-level annotations, both are richly characterized by clinical measurements, including BMI, blood pressure, and glycemic markers.

Another prominent large-scale resource is NHANES [10], which stands out for its population-representative design and its deep metabolic and anthropometric characterization, including

Table 1
Comparison of publicly available wrist-worn and physical activity datasets.

Dataset	Recording Protocol	Sensor Modalities	Device	Sampling Rate	Participant Metadata	Metabolic Information
CAPTURE-24 [7]	24 h	Triaxial accelerometer, wearable camera images	Axivity AX3	100 Hz	Demographics	Physical activity annotations, MET estimations
UK Biobank [8]	7 days	Triaxial accelerometer	Axivity AX3	100 Hz	Demographics, clinical markers, socioeconomic status	Sleep logs, macro-level MET estimates
China Kadoorie Biobank [9]	Up to 7 days	Triaxial accelerometer	Axivity AX3	100 Hz	Demographics, socioeconomic status, anthropometrics, medical history	Sleep, sedentary, and physical activity indicators
NHANES [10]	7 days	Triaxial accelerometer	ActiGraph GT3X+	80 Hz	Demographics, socioeconomic status, comprehensive health profiles	Step counts, wear-time indicators, MET estimates
HUNT4 [11]	Up to 7 days	Triaxial accelerometer (multi-position)	Axivity AX3	50 Hz	Demographics, health questionnaires, blood biochemistry, clinical measurements	None
Proposed: IMU50 [15]	5 days	Triaxial accelerometer, triaxial gyroscope, skin temperature, dual-channel PPG	ActiGraph LEAP	IMU at 128 Hz, PPG at 25 Hz, skin temperature every minute, METs and caloric expenditure every hour	Demographics (age, sex, education level), anthropometrics (BMI, waist/hip circ.), lifestyle (smoking/alcohol)	METs, caloric expenditure, skin temperature, and PPG

lipid panels and insulin levels measured in clinical settings. Moving beyond single-device protocols, the HUNT4 dataset [11] occupies a distinct niche by adopting a multi-sensor configuration with accelerometers worn simultaneously on the thigh and lower back, thereby providing richer biomechanical information regarding posture and movement. HUNT4 participants also underwent comprehensive clinical assessments, including body composition analysis, blood biochemistry, and cardiorespiratory fitness testing, making it exceptionally well-suited for investigations linking learned activity representations to detailed physiological phenotypes.

Despite their massive scale and clinical depth, these existing datasets share a common technical limitation: they rely exclusively on triaxial accelerometry, completely lacking concurrent gyroscope, skin temperature, or photoplethysmography (PPG) measurements. Furthermore, while some offer rich cross-sectional clinical profiles, they lack continuous or high-resolution metabolic references such as hourly energy expenditure captured under free-living conditions. Consequently, they are suboptimal for tasks requiring multimodal signal modeling, multi-sensor fusion, or the study of inter-individual variability linked to body composition.

The present dataset directly addresses these gaps by providing continuous, multi-day wrist IMU recordings that encompass both triaxial accelerometer and triaxial gyroscope signals sampled at a high frequency of 128 Hz. Collected from 50 participants over more than five consecutive days and nights under unconstrained free-living conditions, these inertial data are synchronized with continuous skin temperature measurements (sampled every minute) and dual-

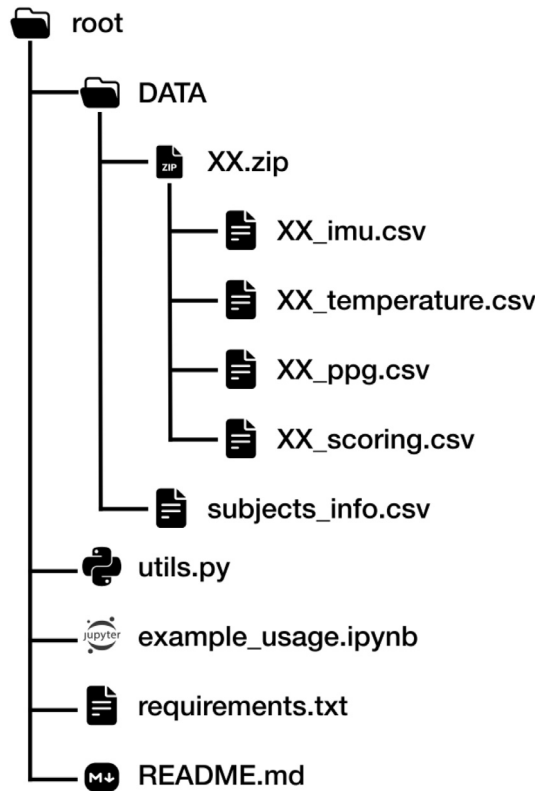


Fig. 1. Structure of the files and folders present in the repository.

channel PPG data (sampled at 25 Hz). Crucially, to enable metabolic outcome modeling, the dataset includes hourly metabolic equivalent of task (MET) and caloric expenditure estimates derived from the established Freedson et al. (1998) algorithms [2]. This physiological profile is further complemented by granular participant-level demographic (age, sex, education), anthropometric (body weight, height, circumference of the hips and waist, BMI), and lifestyle factors (alcohol and tobacco consumption). The combination of high-resolution multimodal sensing, extended tracking duration, and continuous metabolic reference data positions this dataset as a unique, complementary benchmark for self-supervised representation learning from unlabeled IMU signals, as well as for supervised regression and classification targeting metabolic and morphological outcomes.

3. Data Description

The dataset presented in this study comprises inertial measurements, skin temperature, MET estimates, caloric expenditure, and PPG signals, along with demographic, anthropometric, and lifestyle information collected from 50 subjects over a 5-day period. The dataset is publicly available on the Zenodo repository [15]. Fig. 1 illustrates the repository's file and folder structure.

All collected data are organized in the DATA directory, which is divided into compressed archives numbered by each subject's ID. The data is compressed in zip format, reducing the total size of the dataset to <50 GB while guaranteeing maximum compatibility across operating systems and devices. Each subject's archive contains four distinct files in Comma-Separated Val-

Table 2
Field descriptions for the *XX_imu.csv* files.

Column Name	Description
Timestamp	ISO 8601 timestamp indicating when the sensor data were collected.
Accelerometer X	Acceleration along the x-axis (g).
Accelerometer Y	Acceleration along the y-axis (g).
Accelerometer Z	Acceleration along the z-axis (g).
Gyroscope X	Angular velocity around the x-axis (°/s).
Gyroscope Y	Angular velocity around the y-axis (°/s).
Gyroscope Z	Angular velocity around the z-axis (°/s).

Table 3
Field descriptions for the *XX_temperature.csv* files.

Column Name	Description
Timestamp	ISO 8601 timestamp indicating when the sensor data were collected.
Temperature	Measured skin temperature value (°C).

Table 4
Field descriptions for the *XX_ppg.csv* files.

Column Name	Description
Timestamp	ISO 8601 timestamp indicating when the sensor data were collected.
Ambient_0	Ambient light intensity captured by sensor 0.
Ambient_1	Ambient light intensity captured by sensor 1.
Green_0	Reflected green light intensity captured by sensor 0.
Green_1	Reflected green light intensity captured by sensor 1.

Table 5
Field descriptions for the *XX_scoring.csv* files.

Column Name	Description
Timestamp	ISO 8601 timestamp indicating when the sensor data were collected.
kcal	Estimated energy expenditure (kcal).
METs	Estimated metabolic equivalent of task (MET) value.

ues (CSV) format: *XX_imu.csv*, *XX_scoring.csv*, *XX_temperature.csv*, and *XX_ppg.csv*. These files are prefixed with the subject's unique identifier (XX_) to match the parent directory name.

The *XX_imu.csv* file contains triaxial accelerometer and gyroscope signals sampled at 128 Hz. [Table 2](#) details the attributes and descriptions of each field within this file.

The *XX_temperature.csv* file consists of skin temperature measurements captured at 1-minute intervals by a dedicated internal thermometer. The file structure is described in [Table 3](#).

The *XX_ppg.csv* files comprise PPG signals sampled at 25 Hz simultaneously from two distinct sensors; the corresponding fields are outlined in [Table 4](#). Photoplethysmography data are acquired by emitting light from a green LED onto the skin and measuring the reflected intensity. To mitigate ambient light interference, each sensor records a baseline reading with the LED turned off. This baseline ambient value can subsequently be subtracted from the active LED measurement to perform noise reduction.

Finally, the *XX_scoring.csv* file provides hourly estimations of caloric expenditure and MET values, as described in [Table 5](#).

In addition to the subject-specific folders, the *DATA* directory includes the *subjects_info.csv* file, which contains demographic, anthropometric, and lifestyle data for all subjects. [Table 6](#) details the attributes and descriptions of each field within this file.

To facilitate data access, the repository includes a Python utility module, *utils.py*, which provides dedicated functions for accessing and managing the different files: *load_imu*,

Table 6
Field descriptions for the *subjects_info.csv* file.

Column Name	Description
Subject	Unique identifier (ID) assigned to each subject.
Sex	Boolean variable indicating the subject's sex (0 = woman, 1 = man).
Age	Age of the subject (years).
Education level	Number of years of formal education completed (years).
Height	Height of the subject (cm).
Weight	Weight of the subject (kg).
Hip_circ	Circumference of the hips (cm).
Waist_circ	Circumference of the waist (cm).
BMI	Body Mass Index of the subject (kg/m ²).
Smoke	Boolean indicator of whether the subject smokes cigarettes or tobacco.
Alcohol	Boolean indicator of whether the subject consumes alcohol.

Table 7
Statistics of continuous physiological and metabolic variables.

	Min	Max	Average	Std
Skin temperature (°C)	16.85	41.34	32.18	1.85
Energy expenditure (kcal/h)	0.00	705.89	58.45	56.95
MET/h	1.00	8.37	1.69	0.64

load_temperature, *load_ppg*, *load_scoring*, *load_subjects_info*, and *list_subjects*. The *load_* functions accept four arguments: the path to the DATA directory, a list of columns to extract, a list of target subject IDs, and the number of samples to load per file, and return a single pandas DataFrame containing the aggregated data. The *list_subjects* function is a utility that returns a list of all subjects available in the dataset. Additionally, the repository features a Jupyter Notebook (*example_usage.ipynb*) demonstrating the data loading and preprocessing pipeline, culminating in the training of a basic deep learning model to predict anthropometric values using exclusively IMU data. Finally, a *requirements.txt* file lists all software dependencies necessary to execute the code, while a comprehensive *README.md* file outlines the repository's overall structure and contents.

3.1. Distribution of continuous and demographic variables

Table 7 summarizes the recorded continuous physiological and metabolic variables, presenting the minimum, maximum, mean, and standard deviation for skin temperature, hourly energy expenditure, and hourly MET. Physiologically, the recorded MET values indicate that the participants are predominantly sedentary to light-active, with an overall mean MET of 1.68, which falls within the light-intensity activity threshold. However, the maximum observed value of 8.37, with a peak energy expenditure of 705.89 kcal/h, demonstrates that the dataset also captures episodes of vigorous physical exertion, such as running or jogging.

To further illustrate these distributions, **Fig. 2** displays the frequency histograms for each measured continuous variable. The distributions of the accelerometer and gyroscope magnitudes corroborate the predominantly sedentary behavior of the subjects. Specifically, the most prominent peaks occur at 1 g for the accelerometer (reflecting gravitational acceleration) and 0 °/s for the gyroscope, which are characteristic of a stationary device. Furthermore, when excluding the gravitational component, the accelerometer distribution exhibits narrow tails that span only between 0.5 g and 1.5 g. This restricted range, combined with the minimal angular velocity captured by the gyroscope, provides quantitative evidence that the subjects engaged in very limited physical movement. This pattern is consistent with the hourly MET and energy expenditure distributions, where the mean values align with basal or light metabolic activity. Lastly, the mean skin temperature of 32.32 °C is consistent with normal peripheral skin temperature profiles typically observed at the wrist in healthy individuals during standard free-living activities.

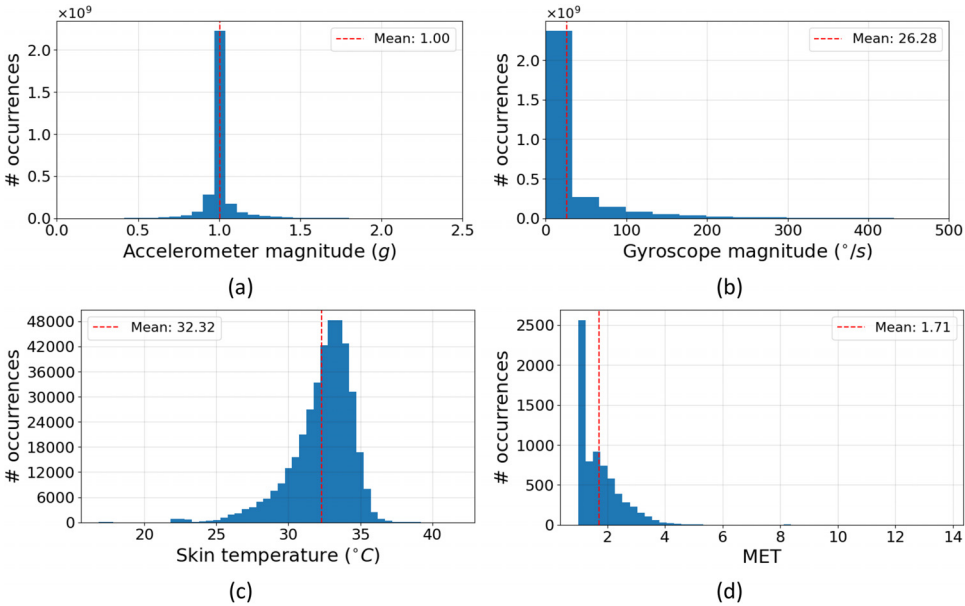


Fig. 2. Frequency histograms illustrating the distribution of continuous variables.

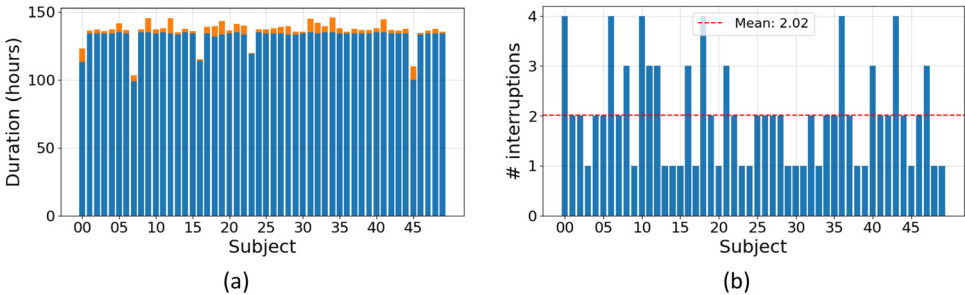


Fig. 3. Stacked plot showing the total recording duration and the cumulative interruption duration (a), and bar plot showing the number of individual data interruptions per subject (b).

3.2. Duration and interruptions of the recordings

To provide a more detailed overview of the recorded signals, Fig. 3 displays bar plots illustrating the total duration of the IMU recording for each subject, the cumulative duration of data interruptions within each recording, and the number of individual interruptions per subject.

As observed, most recordings span approximately 136 h (equivalent to 5.6 days). However, four subjects exhibit shorter durations of around 100 h. Regarding the data interruptions, the average cumulative duration per subject is approximately 4.3 h. However, this value varies significantly across subjects, ranging from over 10 h for some participants to around 2 h for others.

The subject cohort averages approximately two gaps per subject, with a maximum of four gaps observed in some individuals. Normalized over the tracking period, this corresponds to an average of approximately 0.35 interruptions per day. These intervals, together with variations in overall recording length, are directly related to the data collection protocol, which required participants to temporarily stop recording to recharge the device.

4. Experimental Design, Materials and Methods

This section provides a detailed description of the subjects included in the IMU50 dataset, the wearable device used for data collection, the data collection procedure, and the software used to organize, process, and assemble the final dataset.

4.1. Subjects

The dataset comprises 50 healthy volunteers (24 women and 26 men); nine participants smoke cigarettes, and seven consume alcohol. Table 8 presents details on the subjects, including sex, age, height, and weight, hip and waist circumference, and BMI, along with education level and alcohol and cigarette consumption. The subjects were anonymized by randomly sorting them and assigning a progressive numerical ID.

Table 9 presents the anthropometric and cumulative physiological profiles of the participants, reporting the minimum, maximum, mean, and standard deviation for each variable. The age range was selected to recruit a heterogeneous sample of mature adults while minimizing the likelihood of severe health complications typically observed in much older populations. In terms of body composition, the subjects exhibit a diverse distribution of heights and weights.

Finally, Fig. 4 shows the frequency distributions of the participants' characteristics. All variables are evenly distributed, with weight peaking around 80 kg, BMI around 27 Kg/m², and hip and waist circumference around 95 cm. Furthermore, the number of years of education tends to be skewed toward the higher end, with most people ranging from 14 to 18 years of education, suggesting a high level of educational attainment.

4.2. Measurement device

Data collection was conducted using six ActiGraph LEAP [1], a medical-grade wrist-worn smartwatch created by Ametris LLC, engineered specifically for clinical trials and scientific research involving human cohorts. Unlike consumer-grade wearables, the LEAP is explicitly designed to meet regulatory standards for medical monitoring and provide raw, unadulterated sensor telemetry under unconstrained free-living conditions.

From a hardware architecture perspective, the LEAP is powered by a low-power, high-performance microcontroller unit (MCU) optimized for continuous real-time signal processing and efficient power management. Local data logging is supported by 1 GB of on-board non-volatile Flash storage, while external interaction and data transfer are enabled via Bluetooth LE 5 and wired USB interfaces. Designed with a compact form factor measuring $3.9 \times 3.9 \times 1.2$ cm and weighing just 24 g, the internal electronics are housed in an IPX7-rated water-resistant enclosure, capable of withstanding immersion up to 1 m for 30 min and powered by a rechargeable lithium-polymer battery designed to support multiple days of continuous multi-sensor recording.

Regarding motion tracking, the device features a dual-accelerometer setup. The primary accelerometer operates at 32 Hz with a fixed dynamic range of $\pm 8g$. The secondary accelerometer, integrated into the high-resolution IMU device, operates at 128 Hz with a configurable dynamic range spanning from $\pm 2g$ to $\pm 16g$. For this dataset, signals were recorded from the secondary accelerometer to leverage its higher sampling rate, while setting its dynamic range to $\pm 2g$, which represents the standard operational sensitivity for capturing human daily physical activity and micro-movements. The IMU also includes a triaxial gyroscope sampled at 128 Hz with a fixed full-scale range of $\pm 2000^\circ/s$.

Photoplethysmography is captured via a dual-channel PPG sensor operating at 25 Hz, utilizing green light at a wavelength of 530 nm to ensure robust peripheral pulse and hemodynamic tracking. Skin temperature is recorded via an integrated digital thermometer operating across a

Table 8
Details of the subjects and their anthropometric and physiological profiles.

Subject	Sex (0 female, 1 male)	Age (years)	Height (cm)	Weight (Kg)	Hip_circ (cm)	Waist_circ (cm)	BMI (Kg/m ²)	Education (years)	Smoke (0 No, 1 Yes)	Alcohol (0 No, 1 Yes)
00	1	50	170	79	96	94	27.3	13	0	0
01	1	58	173	65	91	78	21.7	18	0	0
02	0	75	156	74.5	118	97	30.6	13	0	0
03	0	60	162	80	110	99	30.5	8	0	0
04	1	47	170	72	100	97	24.9	18	0	0
05	1	71	170	73	98	91	25.3	11	0	0
06	0	67	152	64	100	91	27.7	11	0	0
07	1	41	184	94	103	97	27.8	18	0	0
08	0	65	152	74	110	101	32.0	16	0	0
09	0	60	163	73	108	101	27.5	8	0	0
10	0	50	158	52	91	77	20.8	18	0	0
11	1	54	163	100	126	119	37.6	10	1	0
12	1	56	176	98	116	103	31.6	18	0	1
13	0	44	157	50	92	71	20.3	13	1	0
14	0	55	164	93	130	123	34.6	8	1	0
15	1	66	178	87	107	98	27.5	18	0	0
16	0	60	165	80	111	92	29.4	18	0	0
17	1	67	170	80	99	101	27.7	13	0	0
18	0	58	160	62	96	79	24.2	13	0	0
19	1	67	178	93	115	118	29.4	13	0	0
20	0	70	158	80	109	102	32	13	0	0
21	1	64	175	90	111	107	29.4	18	1	0
22	1	59	197	96	109	101	24.7	13	0	0
23	1	55	183	75	104.5	90	22.4	18	0	0
24	1	68	160	66	97	95	25.8	13	0	0
25	1	62	176	83	102	100	26.8	18	0	0
26	0	40	167	57	92	76	20.1	18	1	0
27	1	60	180	90	110	107	27.8	13	0	1
28	1	66	175	66	95	83	21.6	18	0	0
29	1	71	180	100	120	122	30.9	13	0	0
30	0	62	155	60	101	88	25.0	13	0	0
31	0	44	158	68	113	99	27.2	13	0	0
32	1	70	180	80	102	89	24.7	8	1	1
33	0	69	155	104	139	126	43.3	5	0	0
34	1	67	174	85	97	101	28.1	13	1	1
35	0	49	160	50	92	64	19.5	18	0	0
36	0	61	167	64	98	80	22.9	18	0	0
37	1	75	164	79	113	112	29.4	13	0	0
38	1	46	185	92	111	101	26.9	18	0	0
39	0	53	168	57	97	77	20.2	18	0	0
40	0	65	160	58	90	86	22.7	13	0	1
41	1	70	174	73	91	89	24.1	18	0	1
42	0	44	165	52	91	71	19.1	18	1	0
43	0	44	173	82	120	90	27.4	13	0	0
44	1	75	173	67	97	94	22.4	17	0	0
45	0	51	150	82	119	110	36.4	18	0	0
46	1	55	185	79	89	96	23.1	18	1	1
47	0	53	162	85	115	110	32.4	13	0	0
48	0	54	162	71.6	102	95	27.3	10	0	0
49	1	63	183	85	104	98	25.4	18	0	0

measurement range of $-20\text{ }^{\circ}\text{C}$ to $+50\text{ }^{\circ}\text{C}$ (sampled once per minute). Additional ambient telemetry is provided by a barometric pressure sensor and an acoustic microphone.

Beyond raw signal acquisition, the platform leverages validated onboard algorithmic pipelines to extract higher-level physiological and behavioral markers. These derived metrics encompass vital signs (i.e., heart rate, heart rate variability, and oxygen saturation), physical activity behav-

Table 9

Cumulative descriptive statistics of the dataset.

	Min	Max	Average	Std
Age (years)	40.00	75.00	59.12	9.58
Education (years)	5.00	18.00	14.48	3.59
Height (cm)	150	197	168.50	10.44
Weight (kg)	50	104	76.54	14.40
BMI (kg/m2)	19.10	43.30	26.95	4.91
Hip circumference (cm)	89	139	104.95	11.28
Waist circumference (cm)	64	126	95.72	13.80

ior (such as step counts, energy expenditure, and time spent in moderate-to-vigorous physical activity), and comprehensive sleep architecture indicators (including total sleep time, sleep efficiency, and REM/non-REM stage classification). The hourly caloric expenditure and MET were automatically computed by the ActiGraph device from the primary accelerometer using the established Freedson algorithms.

To account for the high computational and energy demands of continuous high-frequency logging across all sensors, device battery autonomy was systematically evaluated. Empirical testing under full recording conditions demonstrated a continuous battery life of approximately 2.7 days (roughly 65 h). In light of these findings, participants were advised to recharge the device every two days to minimize data interruption over the monitoring period.

4.3. Software

Raw sensor data collected by the devices were extracted using the ActiLife 7 software suite [12]. ActiLife serves as the central software platform for device initialization, data offloading, and quality control across ActiGraph hardware. It enables the conversion and export of proprietary raw binary files into accessible, interoperable formats, such as CSV files, while ensuring precise timestamp synchronization.

Following initial extraction and conversion via ActiLife, all data processing, dataset curation, feature extraction, and graphical plots were executed within a Python environment on a dedicated high-performance workstation. The data signals were loaded, cleaned, and aggregated into the final structured dataset files using the pandas library [13], while all figures presented throughout this paper were generated using matplotlib [14].

4.4. Collection procedure

Participants aged 40 to 75 years were recruited through public advertisements targeting the geographic areas surrounding the provinces of Chieti and Pescara, within the Abruzzo region of Italy. To be eligible for inclusion, participants were required to have no self-reported history of motor impairments or neurological, cardiovascular, and respiratory diseases. Participant selection, monitoring, and all the administrative procedures were conducted at the Department of Neuroscience, Imaging and Clinical Sciences of the University “G. d’Annunzio” of Chieti-Pescara.

Each participant attended an in-person meeting at the university laboratory. During this visit, a member of the research team briefed the participant on the study’s scientific objectives, thoroughly explained device operation, and obtained written informed consent. Participants also completed a structured questionnaire to collect personal and health metadata to establish their profile. Following onboarding, each participant was issued one measurement device, complete with its charging dock and packaging, with instructions to wear it on their non-dominant wrist for a minimum of five consecutive days and nights.

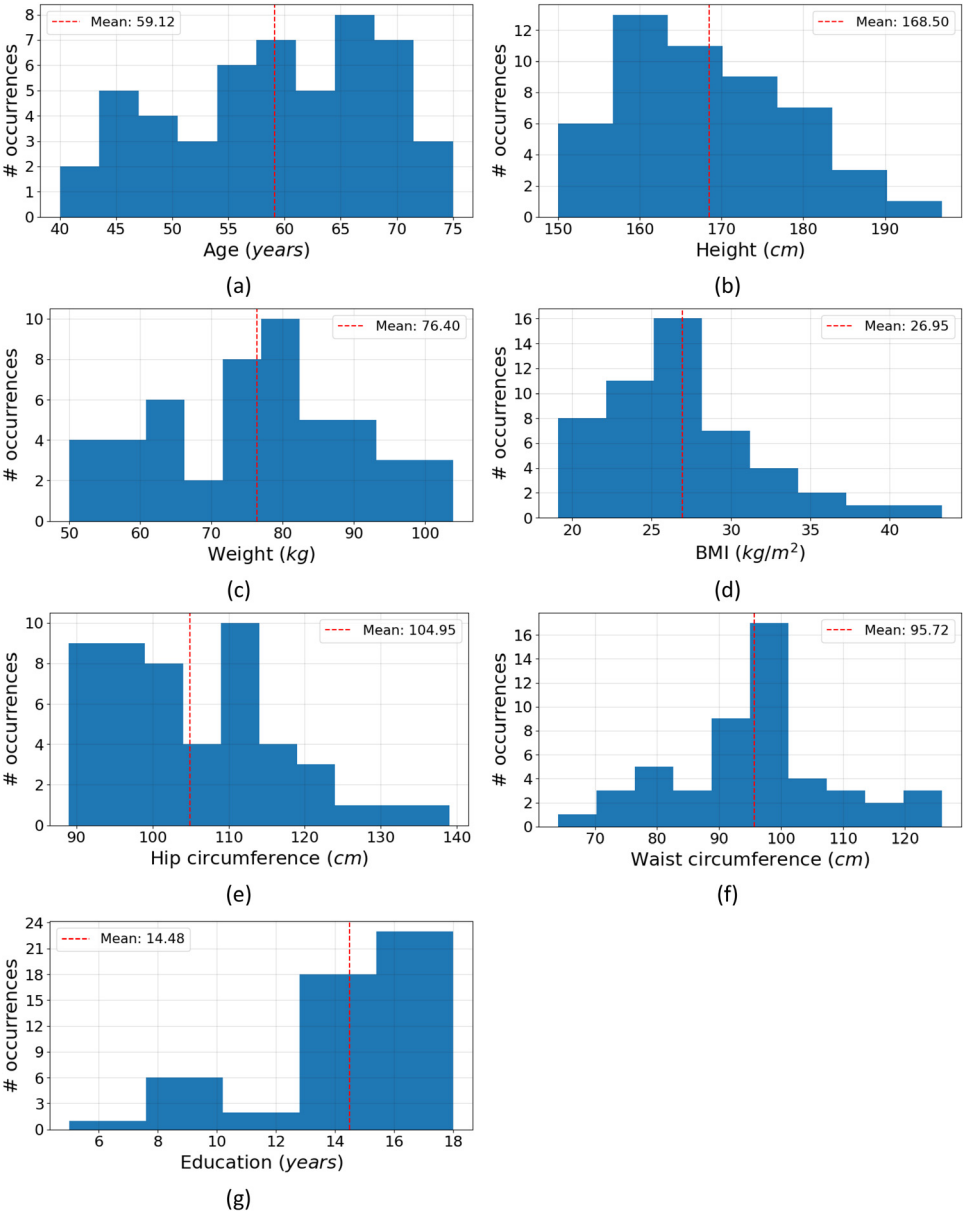


Fig. 4. Frequency histograms illustrating the distribution of participants' demographic characteristics.

Initially, 65 participants were recruited, of which 50 remained after disregarding subjects with recordings exhibiting severe data loss or signal corruption. Participants wore the device continuously throughout their daily routines. The acquired data encompass activities performed within private homes as well as various indoor and outdoor public environments. Continuous wear was required across all conditions, including during sleep and while showering. To capture authentic daily living behaviors, participants were explicitly instructed to maintain their normal routines without altering their lifestyle due to the device. Regarding battery management, par-

ticipants were asked to minimize the number of charging sessions while preventing complete battery depletion. When recharging was necessary, they briefly paused recording via the device interface and resumed continuous acquisition immediately upon re-donning the device. Upon completing the experiment, participants returned the hardware to the laboratory.

Limitations

The dataset does not include ground-truth activity labels, which limits its use for supervised activity recognition benchmarking without additional annotation effort. Furthermore, although participants were instructed to maintain their normal daily routines without restriction, no location data or contextual labels were collected during acquisition. Consequently, the exact proportion of in-home versus out-of-home recording time cannot be quantified. Metabolic estimates are derived from an algorithm validated by Freedson et al. in 1998 [2], whose internal parameters and implementation are not publicly disclosed. The hourly temporal resolution of MET values precludes fine-grained energy expenditure labelling at the window level. The cohort is limited to 50 participants recruited at a single institution, which may not capture the full diversity of body composition, physical fitness levels, and cultural activity patterns present in the general population. Skin temperature data were not independently validated against clinical thermometry. PPG signals are collected continuously throughout the day without removing artifacts due to movement or possible incorrect positioning of the smartwatch.

Ethics Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Biomedical Research Ethics Committee of “C.Et.R.A.—Comitato Etico Regione Abruzzo” (protocol code 1782 of 8 May 2023). Participants aged 40 to 75 years were recruited through public advertisements. Eligible participants had no history of motor impairment, neurological, cardiovascular, and respiratory diseases, and provided written informed consent.

CRedit Author Statement

Irene La Fratta: Investigation, Formal Analysis, Software, Visualization, Writing-Original draft, Writing- Reviewing and Editing. **Lorenzo Calisti:** Methodology, Software, Validation, Data curation, Writing-Original draft, Writing-Reviewing and Editing, Visualization. **Leonardo Bigelli:** Data curation, Writing-Original draft, Writing-Reviewing and Editing. **Piero Chiacchiaretta:** Writing-Reviewing and Editing. **Alessandra Mascitelli:** Writing- Reviewing and Editing. **Damiano D'Ardes:** Writing-Reviewing and Editing. **Piero Di Carlo:** Funding acquisition, Writing-Reviewing and Editing. **Francesco Lucertini:** Writing-Reviewing and Editing. **Antonio Ferretti:** Funding acquisition, Supervision, Writing-Reviewing and Editing. **Emanuele Lattanzi:** Conceptualization, Methodology, Data curation, Writing-Original draft, Writing-Reviewing and Editing. **Raffaella Franciotti:** Conceptualization, Funding acquisition, Methodology, Project administration, Writing-Reviewing and Editing.

Data Availability

IMU50: a high-frequency multi-modal wrist-worn IMU dataset of 50 subjects over 5 continuous days for supervised and self-supervised learning in free-living conditions (Original data) (Zenodo)

Acknowledgements

This research was funded by the [European Union](#), [NextGenerationEU](#), under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 (M4C2), Investment 1.5 Call for tender No 3277 of 30 December 2021, Italian Ministry of University, Award Number ECS00000041, Project Title “Innovation, digitalization, and sustainability for the diffused economy in Central Italy”, Concession Degree No 1057 of 23 June 2022 adopted by the Italian Ministry of University. CUP: D73C22000840006 (University “G. d’Annunzio” of Chieti-Pescara, Italy), CUP: H33C22000430006 (University of Urbino, Italy)

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] ActiGraph LLC. ActiGraph LEAP. Accessed: 22-07-2026. <https://ametriz.com/actigraph-leap>
- [2] P.S. Freedson, E. Melanson, J. Sirard, Calibration of the computer science and applications, Inc. Accelerometer, *Med. Sci. Sports Exerc.* 30 (5) (1998) 777–781 PMID: 9588623, doi:[10.1097/00005768-199805000-00021](https://doi.org/10.1097/00005768-199805000-00021).
- [3] R.P. Troiano, D. Berrigan, K.W. Dodd, L.C. Mâsse, T. Tilert, M. McDowell, Physical activity in the United States measured by accelerometer, *Med. Sci. Sports Exerc.* 40 (1) (2008) 181–188.
- [4] J. Staudenmayer, S. He, A. Hickey, J. Sasaki, P. Freedson, Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements, *J. Appl. Physiol.* 119 (4) (2015) 396–403.
- [5] H. Yuan, S. Chan, A.P. Creagh, C. Tong, D.A. Clifton, A. Doherty, Self-supervised learning of accelerometer data provides new insights for sleep and its association with mortality, *NPJ. Digit. Med.* 7 (2024) 86.
- [6] C.Y. Tang, C. Cho, J. Zhang, SelfHAR: improving human activity recognition through self-training with unlabeled data, *Proc. ACM. Interact. Mob. Wearable Ubiquitous. Technol.* 5 (1) (2021) 1–30.
- [7] S. Chan, Y. Hang, C. Tong, A. Acquah, A. Schonfeldt, J. Gershuny, A. Doherty, CAPTURE-24: a large dataset of wrist-worn activity tracker data collected in the wild for human activity recognition, *Sci. Data* 11 (1) (2024) 1135.
- [8] P.M. Matthews, C. Sudlow, The UK Biobank, *Brain* 138 (12) (2015) 3463–3465.
- [9] Yuanyuan Chen, et al., Device-measured movement behaviours in over 20,000 China Kadoorie Biobank participants, *Int. J. Behav. Nutr. Phys. Act.* 20 (2023) 138.
- [10] L. Koffman, J. Muschelli, Minute level step counts and physical activity data from the National Health and Nutrition Examination Survey (NHANES) 2011–2014 (version 1.0.1), *PhysioNet* (2025) RRID:SCR_007345, doi:[10.13026/9n0r-tv02](https://doi.org/10.13026/9n0r-tv02).
- [11] K. Hagen, A.N. Åsberg, B.L. Uhlig, E. Tronvik, E. Brenner, T. Sand, The HUNT4 study: the validity of questionnaire-based diagnoses, *J. Headache Pain.* 20 (1) (2019) 70, doi:[10.1186/s10194-019-1021-0](https://doi.org/10.1186/s10194-019-1021-0).
- [12] ActiGraph LLC. ActiLife 7 User’s manual. Accessed: 21-07-2026. <https://ametriz.com/actilife>
- [13] The pandas development team, pandas-dev/pandas: pandas (Version v3.0.4) [Computer software], Zenodo (2026), doi:[10.5281/zenodo.21003741](https://doi.org/10.5281/zenodo.21003741).
- [14] J.D. Hunter, Matplotlib: a 2D graphics environment, *Comput. Sci. Eng.* 9 (3) (2007) 90–95.
- [15] I. La Fratta, L. Calisti, L. Bigelli, P. Chiacchiaretta, A. Mascitelli, D. D’Ardes, P. Di Carlo, F. Lucertini, A. Ferretti, E. Latanzi, R. Franciotti, IMU50: a high-frequency multi-modal wrist-worn IMU dataset of 50 subjects over 5 continuous days for supervised and self-supervised learning in free-living conditions, Zenodo (2026), doi:[10.5281/zenodo.21468410](https://doi.org/10.5281/zenodo.21468410).