Mechanical Systems and Signal Processing 168 (2022) 108472

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

A framework for occupancy detection and tracking using floor-vibration signals

Slah Drira [a,b,*], Ian F.C. Smith [a,b]

[a] School of Architecture, Civil and Environmental Engineering (ENAC), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
[b] Future Cities Laboratory, Singapore-ETH Centre, Singapore

ARTICLE INFO

Keywords:
Footstep-induced floor-vibrations
Occupant detection
Support vector machine
Occupant tracking
Model-based data-interpretation
Structural behavior
Walking-gait variability

ABSTRACT

In sensed buildings, information related to occupant movement helps optimize important functionalities such as caregiving, energy management, and security enhancement. Typical sensing approaches for occupant tracking rely on mobile devices and cameras. These systems compromise the privacy of building occupants and may affect their behavior. Occupant detection and tracking using floor-vibration measurements that are induced by footsteps is a non-intrusive and inexpensive sensing method. Detecting the presence of occupants on a floor is challenging due to ambient noise that may mask footstep-induced floor vibrations. In addition, spurious events such as door closing and falling objects may produce vibrations that are similar to footstep impacts. These events have to be detected and disregarded. Tracking occupants is complicated due to uncertainties associated with walking styles, walking speed, shoe type, health, and mood. Also, spatial variation in structural behavior of floor slabs adds ambiguity to the task of occupant tracking, which cannot be addressed using data-driven strategies alone. In this paper, a framework for occupant detection and tracking is developed. Occupant detection is carried out based on signal information. This method outperforms existing threshold-based methods. Support-vector-machine classifiers, trained with time and frequency-domain features, successfully distinguish footsteps from spurious events and determine the number of occupants walking simultaneously. A model-based data-interpretation approach is used for occupant tracking. Structural-mechanics models are used to identify a population of possible occupant locations and trajectories. Up to two occupants can be tracked by accommodating systematic bias and uncertainties from sources such as modeling assumptions and variability in walking gaits. A hybrid framework for occupant detection and tracking that combines model-free approaches for occupancy detection with structural behavior models for tracking is developed and tested on two full-scale case studies. These studies successfully validate the utility of the framework for buildings having sparse sensor configurations that measure floor vibrations.

1. Introduction

Recent progress in sensing and computing technology has resulted in reliable and economic sensors and low-cost computing infrastructure to study human-building interactions. Available technologies have encouraged the development of an automated understanding of occupant information in smart buildings with the goal to optimize important functionalities such as security

* Corresponding author.
E-mail address: slah.drira@gmail.com (S. Drira).

https://doi.org/10.1016/j.ymssp.2021.108472
Received 5 February 2021; Received in revised form 13 August 2021; Accepted 20 September 2021
Available online 30 December 2021

enhancement [1], healthcare [2,3], as well as space and energy management [4–7].

Proposals for sensing technologies that detect occupants inside buildings have included optical sensors [8–10], and radio-frequency devices [11–13]. Occupant detection using radio-frequency devices have been shown to depend on a highly instrumented infrastructure [14–16] that needs regular maintenance. Although detection of occupants using radio-frequency devices has improved in accuracy over recent years [13,17], accuracy is still compromised due to multi-path problems [18] that are induced by elements such as walls and furniture [14,18–20] in a sparsely instrumented infrastructure. Also, wearable-device-based approaches are intrusive and require occupants to carry permanently connected devices. This has led to problems when a device was either off, not connected to the network or when the user decided not to wear a device [21]. These studies have resulted in inaccurate detection and localization of occupants.

Video recording and to a lesser extent, motion sensors, undermine the privacy of indoor occupants due to their intrusive nature [22–24]. For instance, cameras in office environments may influence the behavior of occupants. Employing optical sensors for occupant detection and tracking requires highly instrumented zones and dense sensor configurations to maintain clear lines of sight and large angles of coverage [25–27]. This has led to dense deployment and high maintenance resulting in impractical and costly infrastructure.

Non-intrusive sensing technologies, such as acoustic sensors [28,29], CO_2 sensors [30,31], and smart-flooring systems [32,33] have been proposed. However, acoustic-based methods have been found to be sensitive to ambient audible noise [28,29,34]. The major limitations of CO_2-based approaches are related to the slow spreading of CO_2 within an indoor space where air ventilation influences the concentration of CO_2, leading to ambiguous interpretations of occupancy levels [34,35]. Smart flooring systems require highly instrumented floors (thousands of sensors) [32,33]. Such systems are not suitable for large full-scale applications. In this study, structural-vibration sensors [36–39] are used since they are non-intrusive and have the potential to avoid the shortcomings discussed above.

Detection of events (from footsteps and other sources) has been carried out by assuming the ambient vibrations (i.e. white noise) has the statistical form of an independent zero-mean Gaussian distribution [21]. Events have been detected as anomalies when vibration amplitudes exceed a previously defined baseline level of ambient vibrations (three standard deviations) [22,40]. Another threshold-based method, based on evaluating the auto-correlation of measured vibrations [41,42] has been used to detect and extract footstep-event signals. However, these approaches show limitations due to surrounding noise that hide vibration events having low signal to noise ratios (SNR) [22].

In order to distinguish footstep events from other events, footstep events were classified using one-class support vector machine (SVM) [43] based on 13 features in time and frequency domains [44,45]. Spurious events such as dropping objects, closing doors and drawers, hitting tables and jumping were included to estimate performance compared with Gaussian process and *k*-nearest neighbors (KNN) classifiers [46]. One-class SVM led to an F1 score of 92 % for footstep classification compared with 38 % for fall classification. However, the type II error, defined as the rate of non-footstep events identified as footstep events, was approximately 10 %. Thus, training with only footstep events might miss-classify spurious events as footstep events.

Event classification using a binary-SVM classifier has been proposed by Drira et al [47]. Based on frequency and time domain features, the classifier was trained by footstep events and several spurious events such as book dropping, opening/closing doors and chair dragging. This study involved the selection of frequency ranges that enhance the time domain features to distinguish between classes. Despite high classification performance (accuracy exceeding 97 %), the classifier has been trained only with footstep events from individual occupants [47].

Detected footstep-event signals have been used to estimate the number of occupants on floor slabs [48–50]. A neural network classifier (ANN) [51] has been proposed to determine whether one or two occupants were walking on several floor slabs [42]. Features that were used to train the ANN classifier were signal periodicity using autocorrelation function, signal energy and cross-correlation between adjacent sensor responses. The proposed ANN classifier has been reported to achieve an average accuracy of 86 %. However, several succeeding footstep-event signals are required for training, since measured footstep-induced floor vibrations are segmented into signals of 5 s duration.

Cross-correlation coefficients between event signals at all sensor locations from each footstep event, σ values of event signals recorded at each sensor location and maximum CPSD of all sensors have been used to train the SVM classifier to determine the number of occupants walking together on a floor, Drira et al [47]. These features have been shown to provide high classification performance with an accuracy exceeding 94 %. However, this classification was limited to two occupants walking together [47].

Most existing studies for occupant localization have been based on data-driven approaches. A typical model-free approach involves the assessment of time-difference-of-arrivals (TDoAs) between footstep-induced floor vibrations at multiple sensors to provide estimations of occupant locations [22,40,45,52]. However, these techniques require a monotonic relationship between signal characteristics and the distance from footstep-impact to sensor locations [21]. They have failed to provide accurate localization for varying-rigidity floors (upper floors) and in the presence of obstructions [40,53]. In these floors, a monotonic relationship between event signals and the distance to sensor locations is not satisfied [54–56].

Although model-free approaches have been shown to provide precise results on slabs on grade, complex structural configurations (upper floors) and the presence of obstructions (such as beams, walls and furniture) limit their applicability [40,53]. TDoA techniques rely on Lamb-wave propagation velocity in a medium to estimate the location of an impact source [40]. The propagation velocity depends on material properties of floor slabs such as modulus of elasticity, density and Poisson's ratio. These properties may vary across the floor. Columns, beams and walls contribute to the heterogeneity of typical floor slabs. These elements result in varying floor rigidities that have affected wave propagation properties which make determination of the propagation velocity a challenging task [57]. This limits the applicability of triangulation-based techniques for occupant localization when floors have varying vertical rigidity

[40].

Mirshekari et al [53] have proposed a new triangulation-based approach for occupant localization that estimates the obstruction mass by characterizing the wave attenuation rate. This mass estimation was used to approximate the propagation velocities for accurate localization. However, case studies, presenting uniform vertical rigidities, were limited to the placement of obstructions with several mass levels that did not exceed 60 Kg between sensors and footstep-impact locations. In real applications, typical floors are obstructed with structural elements such as columns, beams and walls and non-structural elements such as heavy furniture. Apart from mass levels, these obstructions can be characterized by their dimensions, materials and connections that may affect significantly the structural behavior of the slab, thereby leading to ambiguous measurement interpretation through triangulation [55].

The presence of structural elements such as beams, columns and walls typically lead to highly dispersive mediums. Due to the dispersive nature of floor slabs, the Lamb-wave propagation induced by footstep impacts is degraded because various wave components of the signal have different propagation velocities. This might result in low SNR vibrations due to the attenuation of signal characteristics, which lead to inaccurate localization [40]. Thus, highly instrumented floors have been required to provide accurate localization results [40,44,58] (a sensor per $\sim 2\ m^2$). This limits the application of TDoA techniques to scenarios where many sensors are deployed.

Moreover, most of the data-driven techniques (triangulation-based approaches) that involved processing and analyzing vibration measurements for localization have been conducted for only single occupants [40,59,60] walking on small-floor areas. In full-scale applications, multiple people walk regularly at the same time on the same floor-slab. Resulting floor-vibration measurements include a superposition of the structural responses from multiple occupants walking with their own speed on their respective trajectory. A few studies have attempted to localize multiple occupants [61,62]. For instance, Shi et al [62] have used continuous wavelet transform (CWT) [63] to decompose the floor vibrations from multiple occupants (up to three). Assuming that people rarely walk in perfect synchrony, filtering the floor vibrations at a high-frequency range (50–125 Hz) was shown to provide better delimitations between footstep excitations (peaks) from multiple occupants. The detected peaks are used for localization based on TDoA technique. Despite the accurate localization results, tests have been carried out on a small unobstructed and highly instrumented space (a sensor per $\sim 2\ m^2$). This concentration of sensors made event-signal separation possible and the estimation of TDoA values achievable. Real scenarios presenting wide and varying rigidity floors may limit the applicability of their methodology.

To date, the most common strategy has been to analyze signals in the absence of a structural behavior model. No research has been found on the application of multiple model-based measurement interpretation for multiple occupant localization using footstep-induced vibrations.

Recent studies have proposed frameworks to identify indoor occupants recorded floor-vibrations induced by footsteps [40,44,49]. Poston et al. [49] have proposed a framework for real-time occupancy tracking that involved separating the floor into zones. Mirshekari et al. [40] has presented a framework for occupant localization that started by detecting event signals from floor responses. A classification module is then used to distinguish between footstep and non-footstep events. Footstep events are subsequently used to provide estimates of occupant locations. More recently, a similar framework has been proposed by Clemente et al. [44] to detect footstep-impact signals from spurious events using supervised-learning classification. In addition to footstep-event detection and classification, the framework included the detection of fall events, occupant localization and identity recognition. However, all of these frameworks are operational only for localizing single occupants on highly instrumented floors (a sensor per $\sim 2\ m^2$) [40,45] having uniform vertical rigidity (slabs on grade) [50,58]. Also, none of these studies proposed a framework to detect, count, and track more than one occupant walking simultaneously on large full-scale varying-rigidity floors (upper floors) using sparse sensing (a sensor per $\sim 10\ m^2$ at least).

This paper contains a proposal for a framework for detection and tracking of building occupants using floor vibrations that are measured from sparse velocity-sensor configurations (one sensor per ~ 12 to $75\ m^2$). The framework combines model-free approaches for occupancy detection, with structural behavior models for tracking. The framework brings together results reported in several recent papers by the authors [47,55,64] to create a coherent workflow of methods. The framework is then validated using two full-scale case studies. Occupant detection is carried out using structural information such as fundamental frequencies for supervised learning classifiers based on a support vector machine (SVM). A novel model-based methodology for accurate occupant tracking using error-domain model falsification (EDMF) [65] is proposed to alleviate the limitations of structural-model-free approaches. Accounting for systematic errors and model bias, EDMF combines information related to measured footstep event signals with simulated physics-based models to identify a population of candidate occupant locations [36]. Simulations model the dynamic responses of the floor that are induced by footstep impacts and they include real structural obstructions without introducing artificial masses.

The paper starts with a presentation of the challenges in the interpreting vibration measurements (Section 2). Description of the occupant detection and tracking framework is presented in Section 3. Model-free occupant detection strategies are explained in Section 4. Details of model-based multiple- occupant tracking strategies are described in Section 5. The application of the framework to a full-scale varying-rigidity floor slab (area of $100\ m^2$) is described in Section 6. The utility of the model-based occupant-tracking methodology is evaluated on another full-scale floor slab (area of $950\ m^2$) in Section 7. Finally, in Section 8, results obtained from both case studies are discussed.

2. Challenges interpreting vibration measurements

Detecting the presence of occupants on a floor is challenging due to variations in rigidities of floor slabs and the presence of obstructions such as beams and walls [66]. Moreover, the dispersive nature of floor slabs creates variations in footstep-impact signatures at various floor locations [40]. This may result in footstep-impact events with low signal-to-noise ratios (SNR).

High levels of ambient noise undermine detection of events that have low SNR values [23]. A footstep-event signal that is recorded by a vibration sensor is compared with amplitude of ambient vibration (three standard deviations, 3σ) is presented in Fig. 1. Threshold-based detection using metrics, such as three standard deviation (3σ) limits, shown in Fig. 1 are not able to detect events when signal amplitudes do not exceed the ambient-vibration level.

Floor-vibration measurements that are recorded by sensors can be caused by activities such as footsteps, closing of doors, moving of furniture and other interactions between occupants and the building space. The duration of dynamic structural responses to an impact event (event-signal duration) depends on the type of event and the walking gaits of occupants. Also, the propagation of Lamb waves in a dispersive medium that are induced by footstep impacts results in shape changes of the floor responses at sensors [40]. This behavior leads to distortion in the time-of-arrival of measured footstep-impact signals. This creates ambiguity in structural responses, both in amplitude and duration, which complicate occupant detection and tracking.

Examples of signals that are induced by chair dragging, door closing and footstep events are shown in Fig. 2. For each event shown in Fig. 2, the duration of structural response recorded by the sensor is different. Chair dragging has larger noise and contribution from higher modes. Door closing excites somewhat lower and higher modes than footstep events. Footstep events are damped faster than door closing events.

Footstep impacts of multiple people walking simultaneously may be fully synchronized (no time offset in the floor response), off-synchronized (varying time offset in the floor response), and staggered (floor response from each individual is separated) [50].Off-synchronized footstep events result in overlapping signals from multiple occupants walking simultaneously on a floor, which make occupant detection and tracking challenging.

Walking gaits of occupants (and thus the impact on the floor) are affected by factors such as their anatomy, walking speed, shoe type, health and mood [67–69]. Since the same occupant may produce several walking patterns, there is inherent variability amongst footstep-impact signatures.

An example of footstep-event signals recorded by the same sensor induced by an occupant walking at the same location is presented in Fig. 3. In Fig. 3a and b, the occupant walks at a slow speed and wearing soft (a) and hard (b) soled shoes. In Fig. 3c and d, the occupant walks at a fast speed and wearing soft (c) and hard (d) soled shoes. Change in walking speed and shoe type leads to a significant variability in amplitudes of footstep-event signals from the same occupant.

3. Framework for occupant detection, localization and tracking

Steps that are part of the occupant detection and tracking framework are shown in Fig. 4. In view of the importance of floor-slab characteristics on vibration response as well as the significant walking-gait variability, as shown by Drira et al [55] and Section 2, this framework combines information on structural behavior and uncertainties from multiple sources to achieve accurate and precise detection and tracking of occupants.

The framework starts with model-free approaches for the detection of footstep events and the determination of the number of occupants. These operations involve extracting event signals, from footstep and other spurious events, and then determining those that correspond to footstep events. After this, extracted footstep-event signals are used to count the number of walking occupants on the floor. Strategies that are involved in the model-free occupant detection strategies are discussed in Section 4.

Event detection requires the understanding of structural characteristics such as the fundamental frequencies of the structure, which are estimated using ambient vibrations. Using structural information (frequency ranges that cover the first few vertical modes of the floor slab), event detection and subsequently signal extraction are carried out (see Section 4.1). Then, a supervised learning classifier based on a support vector machine (SVM) is used to distinguish between extracted footstep and non-footstep-event signals (see Section 4.2). The footstep-event signals are then used to count the number of occupants on the floor using another SVM classifier (see Section 4.3).

The number of occupants is used to select appropriate model simulations for occupant localization as shown in Fig. 4. Occupant localization is carried out using a model-based approach for each detected footstep event. Model-based localization involves the combination of information from footstep-event signals with physics-based simulation models. Inferring potential structural behavior models from measurement data to localize occupants is an inverse engineering task. Systematic errors and high levels of biased uncertainty from multiple sources mean that model-based approaches are ill-posed. Wrong behavior predictions may be obtained if single so-called “optimal” solutions are proposed. A model-based measurement interpretation approach that takes multi-source biased uncertainties explicitly into account is required to find sets of locations of occupants. Error-domain model falsification (EDMF) is used

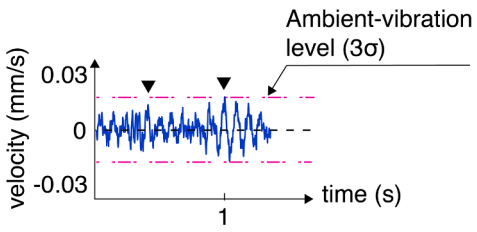

Fig. 1. Footstep-event signals compared with ambient-vibration levels of three standard deviations (3σ) represented by dashed-dotted lines. Triangles locate two footstep events.

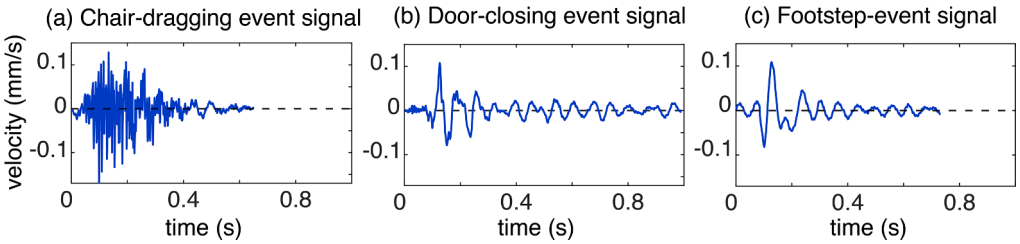

Fig. 2. (a) chair dragging, (b) door closing and (c) footstep event signals. Event signals are recorded by the same sensor.

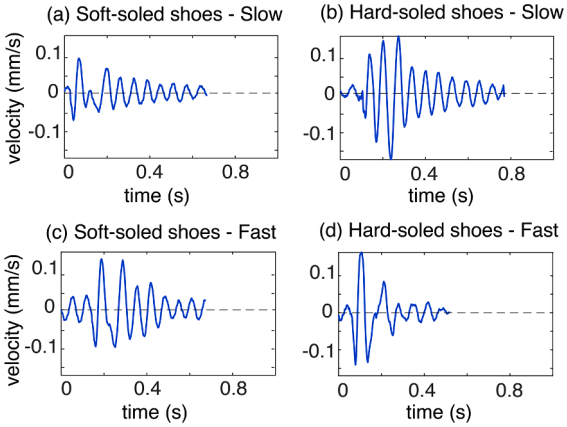

Fig. 3. Footstep-event signals recorded by the same sensor from the same occupant walking at the same location. (a) and (b) the occupant walks with soft-and-hard-soled shoes at a slow speed. (c) and (d) the occupant walks with soft-and-hard-soled shoes at a fast speed.

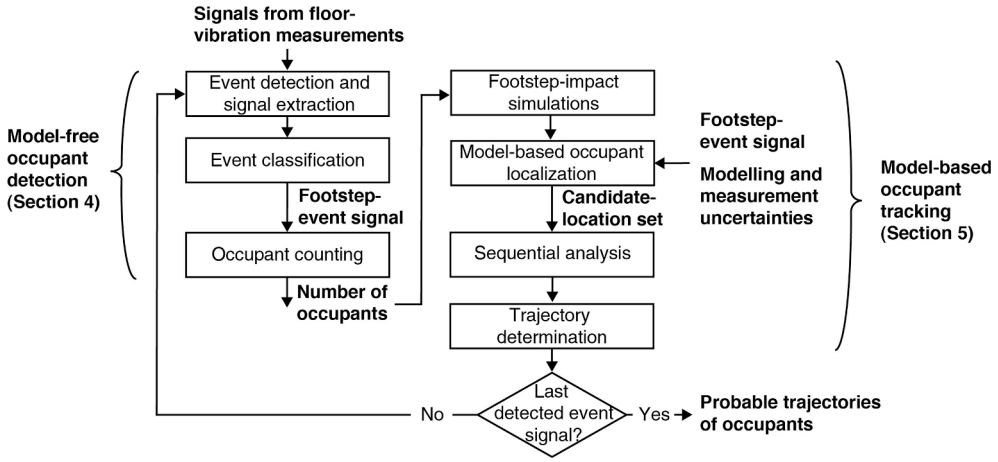

Fig. 4. Occupant detection, localization and tracking framework using footstep-induced floor vibrations.

[36,55,70].

EDMF is a model-based data interpretation approach that is applied to identify a population of possible locations of occupants. In this paper, the population of possible locations of occupants is referred to as a candidate-location set (CLS). Details of model-based occupant localization are described in Section 5.1.

In Fig. 4, the CLS related to each succeeding detected event is then subjected to a sequential analysis based on information from the previously detected event to enhance the precision of occupant location for the current footstep event. Sequential analysis, as explained in Section 5.2, involves the assumption that the distance between two successive impact events cannot exceed a predefined step-length.

Finally, the CLS of each detected event, resulting from sequential analysis, is investigated to determine possible trajectories. The trajectory-determination operation, as explained in Section 5.3, involves the assumption that occupants walk until reaching their

destinations without backtracking. Once CLSs of all detected event signals are investigated, paths connecting remaining candidate departures with possible arrivals that define candidate locations are taken as candidate trajectories.

4. Model-free occupant detection

4.1. Event detection

The first step of the framework involves detecting all possible events from floor vibrations and subsequently, extracting the relevant event signals as described in detail by Drira et al [47]. The event-detection strategy is intended to capture the occurrence times that contains prominent indications of possible events within the measurements.

Occupants strike the floor differently. Their footsteps unequally activate several bending modes of the structure. Information from multiple frequency components of floor vibrations has potential to enhance the detection of events with low SNR values [37,47]. These frequency ranges contain the fundamental bending modes of the structure. Using ambient vibrations, prominent peaks in the first singular values of the cross-power spectral density (CPSD) [71] help delimit the range with most energy contribution. This range is then divided into at least four equivalent and overlapping frequency ranges to cover the fundamental vertical modes of the structure. Floor vibrations are then decomposed using continuous wavelet transform (CWT) [63] and reconstructed using inverse wavelet transform (IWT) at these frequency ranges in order to filter extraneous signal elements.

Filtered floor vibrations are then segmented into windows with a length 0.2 s moving with an increment of 0.1 s. The duration of the moving window helps cover all possible walking frequencies [47,68]. The maximum standard deviation, comparing all sensors, $STD_{max,f}(i)$, at a frequency range, f, as shown in Eq. (1), is used as metric to find abrupt variations in data due to possible footsteps.

$$STD_{max,f}(i) = \max_{s}\left(\sqrt{\frac{1}{N_s-1}\sum_{j=1}^{N_s}\left(S_{i,s,f}(j)-\mu_{i,s,f}\right)^2}\right) \tag{1}$$

In Eq. (1), the signal, $S_{i,s,f}$ that is delimited by the segment number, i, is recorded by the sensor, s, and processed using CWT and IWT at frequency range, f. The number of data defining the signal segment, $S_{i,s,f}$, is denoted as N_s. The mean value of the signal segment, $S_{i,s,f}$, is denoted as $\mu_{i,s,f}$.

Similarly, ambient vibrations (no footstep activity) are decomposed at the same frequency ranges to establish detection thresholds (DT_f). DT_f is defined using STD_{max,f,S_a} values of filtered ambient vibrations (S_a) at frequency range (f) as shown in Eq. (2). For each frequency range, f, the maximum value of the resulting STD_{max,f,S_a} values are taken to be a DT_f. In Eq. (2), i is the signal segment index.

$$DT_f = \max_{i}\left(STD_{max,f,S_a}\right) \tag{2}$$

It has been shown that a local maximum resulting from $STD_{max,f}$ values corresponds to a signal segment that contains prominent magnitudes (from all sensors) of an event signal [47,70]. Thus, local maxima within $STD_{max,f}$ values that exceed DT_f over at least one filtered signal indicate the occurrence times of possible events. Since the minimum time between two footsteps when an occupant walks at maximum speed of 2.5 Hz [68], each local maximum has to be defined within an interval of 0.4 s at least. Moreover, multiple occurrence times pointing to a same event may occur since times of local maxima resulting from $STD_{max,f}$ values at each frequency range are not equivalent. In order to avoid extracting the same event signal twice, the event-occurrence time has to be greater than the ending time of the last extracted event signal.

Starting and ending times of detected events are determined dynamically based on a filtered signal at frequency range greater than the first natural frequency of the structure that has been shown to provide better event delimitation in the time domain [37,47,70]. Thus, $STD_{max,f}$ (Eq. (1)) values that are defined above the natural frequency of the floor slab are used as inputs for event extraction. A backward and forward search of $STD_{max,f}$ values from the local maximum is carried out to determine the starting and ending times for each event. A local maximum is hence compared with its preceding and succeeding $STD_{max,f}$ values to search for two local minima defining the signal segments that contain the starting and ending times of detected events. Subsequently, the sums of the absolute values of the amplitudes of the raw signal of all sensors are computed for the signals delimited by the determined segments. Minimum values of the resulting trends define the starting and the ending times of a detected event. These times also serve to extract the event signal separately for all sensor locations. Signal-extraction operation to determine dynamically the starting and ending times of each detected event signal at sensor locations is then applied as described in detail by Drira et al [47].

4.2. Event classification

Floor vibrations are affected by indoor activities such as chair dragging and falling objects as well as external sources such as traffic and wind. These sources might produce vibration amplitudes that are similar to footstep-impact signals, as explained in Section 2. Thus, the next step of the occupant detection and tracking framework is event classification to distinguish between footstep and non-footstep event signals as described in details by Drira et al [47]. This is carried out using a supervised learning technique based on a support-vector-machine (SVM), which provides good performance with small training sets with respect to feature numbers compared with neural-network-based methods [24,72]. A binary-SVM classifier is trained with footstep and non-footstep events to improve accuracy and prevent miss-classification of spurious impulses when compared with one-class SVM.

Features are assessed in time and frequency domains to discriminate between footstep and non-footstep events [44]. This is due to the natural vibration modes of structures that affect the floor vibrations. Frequency-domain metrics include the frequency value that corresponds to the maximum of the first singular values of CPSD (see Section 4.1) of all sensors (FSV_{max}) and the centroid of first singular values of CPSD (C_{CPSD}). Time-domain metrics include standard deviation (σ) and kurtosis (K_r) of event signals. Other metrics such as maximum difference in amplitudes (Δ_{amp}), root-mean-square (RMS) and median (M_d) are found to be correlated with the σ and therefore excluded from the training process to avoid overfitting.

Since events from various sources may have differentiable signal characteristics at different frequency ranges, time-domain metrics are assessed for filtered event signals at various frequency ranges using CWT. The frequency band characterizing the vibration sensors is divided into equivalent ranges with an overlap to cover signal components with low and high frequency ranges. The size of each frequency range is fixed using engineering judgment [47]. The maximum and the average values of all sensors are evaluated for all time-domain metrics.

Time-domain metrics assessed at specific frequency ranges that maximize the discrepancy between footstep and other events are selected as features for classification. In order to select the appropriate frequency ranges for each time-domain metric, independent statistical testing is used. This is carried out using null-hypothesis based on the Kolmogorov-Smirnov test [73] that estimates the discrepancy level between footstep and non-footstep populations for each time-domain metric and for each frequency range [47]. Thus, for each time-domain metric, the frequency range that has the highest discrepancy level between footstep and non-footstep populations is selected.

4.3. Occupant counting

Footstep-event signals can originate from one or multiple people walking together on the floor slab as presented in Section 2. The next step of the occupant detection and tracking framework is the determination of the number of occupants on the floor. The occupant counting operation is carried out using a multi-class SVM classification as described in detail by Drira et al [47].

The resulting floor vibrations include a superposition of the structural responses from multiple occupants walking at their own speed on their trajectories. In addition, footstep-induced floor vibrations at sensor locations are further altered by the structure and depend on footstep locations [55,70]. Although floor vibrations are influenced by the structure and even though there is often no reliable relationship between amplitudes and proximity, the highest amplitudes at sensor locations are generally recorded when footstep impacts are very close [47,55,70]. Thus, cross-correlation coefficients between event signals at all sensor locations are used as features to train the multi-class SVM classifier. The cross-correlation coefficient matrix is calculated based on the Pearson linear-correlation method [74].

In addition, standard deviation (σ) values of event signals recorded at each sensor location and maximum CPSD of all sensors are used as features to increase the classification performance to determine the number of occupants walking on the floor [47]. It has been shown that σ values and maximum CPSD of event signals are correlated to the impact force induced by footsteps at a sensor [38] and have higher magnitudes induced by multiple occupants than values from single occupants [70].

5. Model-based occupant tracking

The model-based occupant tracking methodology is intended to identify possible locations of either a single occupant or multiple occupants walking simultaneously on a floor slab. This methodology presents the final step of the occupant detection and tracking framework, Section 3. This methodology accounts for uncertainties from multiple sources to achieve accurate and precise tracking of occupants.

Measured and simulated footstep-event signals are inputs of the model-based occupant localization operation. The number of occupants is used to select the appropriate model simulations for occupant localization as shown in Fig. 4. Explicitly accounting for uncertainties from multiple sources using error-domain model falsification (EMDF) as presented in detail by Drira et al [36,55,75], measured footstep-event signals are compared with the selected impact simulations at possible locations. Impact locations that do not contradict measured floor response define a candidate-location set (CLS). This operation is carried out separately for each footstep-event signal. Subsequently, resulting CLSs of event signals induced by a signal occupant or multiple occupants are used to determine possible trajectories. Further details on model-based occupant localization are presented in Section 5.1.

Possible departure/arrival points are defined prior to sequential analysis (Section 5.2) trajectory-determination (Section 5.3) operations as inputs. Candidate locations (CLs) resulting from the first detected event help determine candidate departures (starting locations). Departure points that do not contain CLs within a radius of twice the distance between the two footsteps (i.e. step length) are rejected. Departure points that are not rejected are taken to be candidate departures.

Prior observations reveal that step length is found to vary between approximately 60 cm and 90 cm with respect to the walking-speed level (from slow to fast walking) [68,76,77]. In this paper, an average step length of 75 cm is chosen for model simulations. This step length is used for a sequential analysis to enhance the precision of resulting CLSs (Section 5.2). CLSs associated with possible departures are then used to determine possible trajectories of each occupant (Section 5.3).

5.1. Occupant localization using model-based approach

5.1.1. Single occupant

Floor vibrations that are induced by occupants walking individually are found to be independent from the consecutive footstep event [36,40]. Thus, footstep-impact simulations for a single occupant are carried out separately using a grid of predefined possible locations. The distance between two possible locations is fixed to be equal to the average step length (75 cm, see Section 5). In this paper, these predefined possible locations are referred to as initial location set and they are represented by dots in Fig. 5.

Model simulation results and measurement values are affected by uncertainty from several sources [36,55,75]. Common modeling uncertainties are related to model assumptions, such as idealized boundary conditions, and omissions. Also, unknown model parameters and an idealized footstep-impact load function contribute to the total modeling uncertainty. In the context of occupant localization, the natural variability in walking gaits significantly contributes to the variability in floor-vibration measurements [55] (see Section 2). These variations are influenced by occupant anatomy, walking speed, shoe type, health and mood [55]. Modelling and measurement uncertainties are estimated based on previous work, engineering judgment and prior floor-vibration measurements [36,55,75].

For occupant-localization applications, parameters to be identified (θ) as shown in Eq. (3), are x and $y$coordinates of possible occupant locations on the floor slab as shown in Fig. 5. Using EDMF, location instances are falsified when residuals between measured ($m_{l,e}$) and simulated ($g_l(\theta)$) responses lie outside the predefined thresholds that are represented by dashed lines in Fig. 5. Localization thresholds, denoted as $T_{high,l,e}$ and $T_{low,l,e}$ in Eq. (3), are determined from the combined uncertainty distribution ($U_{c,l}$ in Fig. 5) and a target reliability of localization that is fixed at 95 %.

$$T_{low,l,e} \leq g_l(\theta) - m_{l,e} \leq T_{high,l,e} \forall l \in \{1, .., n_m\} \tag{3}$$

Given an initial location set, for each detected footstep event (e), all location instances whose residuals lie inside the localization thresholds ($T_{low,l,e}$ and $T_{high,l,e}$) at each sensor location (l) are accepted and form the candidate-location set (CLS) as represented by the encapsulated dots in Fig. 5. Together, these model instances are the CLS that includes the real location of a footstep event (see real footstep location in Fig. 5). Candidate locations (CLs) are generated for each detected footstep event.

The identification of position vector θ (x and y coordinates) of occupants is not computationally expensive since measured footstep-event signals are compared with pre-simulated vibrations from footstep impacts at a grid of possible locations. This allows repeated use of the same simulation results for multiple comparisons with all measured footstep-event signals. Model falsification thus enables a near-real-time localization of occupant footsteps. EDMF provides a population of possible locations. All CLs are assumed to be equally probable due to the lack of information of the true uncertainty distributions. Finally, since EDMF is a constrained satisfaction procedure, there are no difficulties associated with convergence as there may be with conventional optimization methods. Superposition of simulations also result in low computational costs for more than one occupant. This is described in the next section.

5.1.2. Two occupants

In full-scale applications, multiple people walk regularly at the same time on the same floor-slab. The resulting floor-vibration measurements include a superposition of the structural responses from multiple occupants walking with their own speed on their

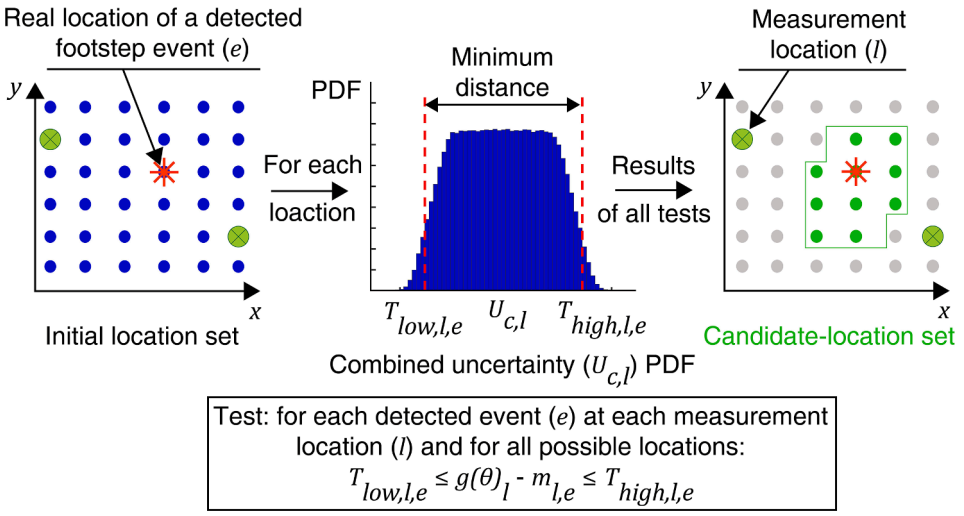

Fig. 5. Occupant localization using error-domain model-falsification (EDMF) starts with an initial location set. Model simulations are compared with measured response in order to identify candidate locations among the initial population. Threshold boundaries are derived from combined uncertainties. Model instances are falsified when the residual value between simulations and measurements exceeds thresholds at any sensor location.

respective locations. Footstep impacts of multiple people may be either fully synchronized, off-synchronized or staggered, as described in Section 2.

Unlike military marching, the footstep impacts of multiple people walking simultaneously are usually off-synchronized. This results in overlapping signals with varying time-offsets in floor responses. To a lesser extent, footstep impacts from multiple occupants may also be staggered, leading to separate floor responses from each individual.

A graphical representation of floor responses resulting from two occupants (O1 and O2) walking together on a floor slab is presented in Fig. 6. In Fig. 6, dots represent the footstep-impact locations related to the two occupants. The circled signals on top of Fig. 6 represent floor responses that are induced by each occupant walking separately at a same speed level (i.e. walking frequency). In this representation, both signals are assumed to have a similar duration (D = 1/walking frequency).

In Fig. 6a and b, two scenarios, in which footstep impacts from occupants O1 and O2 are off-synchronized are presented. The resulting floor responses recorded at the sensor location (S) are generated considering two time-offsets: one-third (a) and two-thirds (b) of the impact duration (D). In Fig. 6c, the scenario when the two footstep impacts are staggered is presented. The floor responses induced by off-synchronized footstep impacts as shown in Fig. 6a and b results in overlapping signals. However, staggered footstep impacts result in a succession of impact signals that are contributed by each occupant.

A similar strategy that is used to localize single occupants, as explained in Section 5.1.1, EDMF is carried out to identify possible locations of two occupants walking simultaneously on a floor slab. Using EDMF, simulated and measured floor responses induced by the two occupants are compared. Footstep-impact simulations for two occupants are generated following three steps.

Initially, the footstep impacts of a single occupant are simulated using a grid of predefined possible locations using a finite-element model. A moderate walking speed of 1.6 Hz (impact duration of 0.625 s) is maintained during simulations since the variability in gaits due to change in speed level during walking tests is accounted for the combined uncertainty for localization (see Section 5.1).

Subsequently, using the initial location set, all possible location combinations, resulting from two occupants, are generated. Then, two time-offsets: one-third (0.2 s) and two-thirds (0.4 s) of a footstep-impact duration are used to superpose contributions from each occupant at each possible location, see Fig. 6a and b. These time offsets are chosen based on prior observations of measured vibrations from two occupants walking together. A fixed duration of 0.625 s of the superposed signals is investigated for localization. Moreover, simulations from a single occupant are included in the initial model instances to account for staggered footstep events, see Fig. 6c.

5.2. *Sequential analysis*

Occupant-localization using model falsification (Section 5.1) may be enhanced by a sequential analysis as presented in Fig. 4. Sequential analysis for footstep-impact localization involves the assumption that a person walks continuously with a fixed step length. This behavior is tested for all CLSs related to each succeeding detected footstep-event based on information from the previously detected event.

A graphical explanation of the sequential analysis is presented in Fig. 7. The size of the CLS at each detected footstep event is further

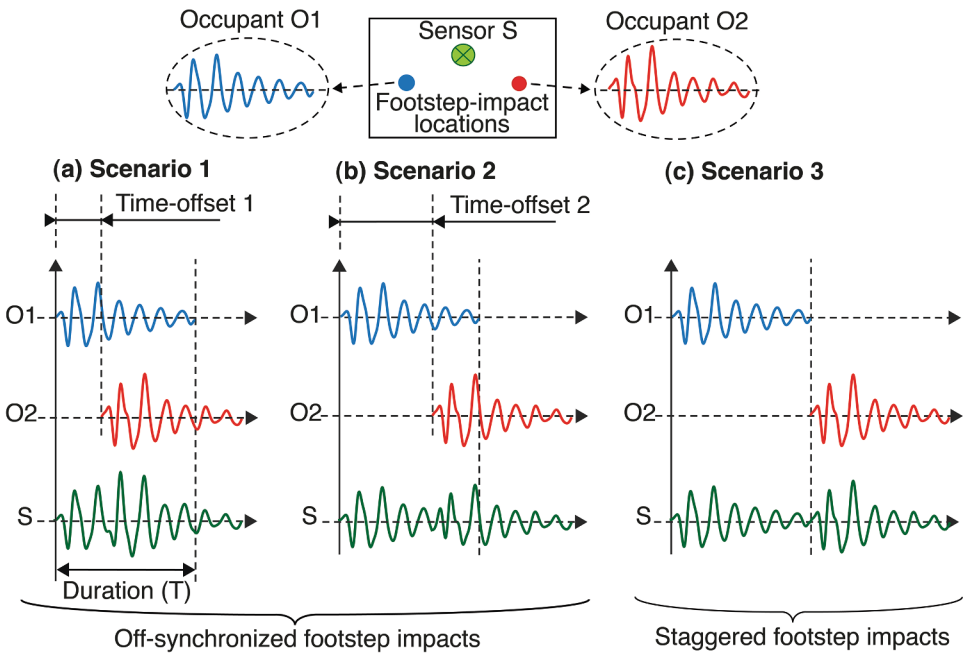

Fig. 6. Floor vibrations resulting from two occupants walking together. (a and b) two scenarios when footstep impacts are off-synchronized. The resulting floor responses recorded at sensor location regarding two time-offsets: one-third (a) and two-thirds (b) of a footstep-impact duration. (c) scenario when footstep impacts are staggered.

reduced using information from the previously detected event. In Fig. 7, each CL of current detected event (s), represented by squares, is compared with all CLs of the previously detected event ($s-1$) that are represented by dots.

When the minimum distance between a CL of current detected event (CL_s) and all CLs of a previously detected event (CLS_{s-1}) is higher than the pre-defined distance between two footsteps, the CL_s is rejected as represented by diamonds in Fig. 7. Assuming that a person carries on walking until reaching their destination, information from previous footstep events helps reduce the size of the CLSs.

The sequential analysis is repeated separately, for each CLS that is allocated to each candidate departure, from the previously detected footstep event as described in Section 5. This operation is mathematically expressed by the following equation:

$$\min\left(\|CL_s - CLS_{d,s-1}\|_2\right). > \text{Steplength}, \forall d \in \{1, .., n_d\} \tag{4}$$

In Eq. (4), A CL of a current detected event (CL_s) is rejected when its minimum distance to all CLs related to a candidate departure of a preceding event ($CLS_{d,s-1}$) exceed the predefined step length.n_d denotes the number of remaining candidate departure from a previously detected event. CLS of each footstep event related to each candidate departure is subsequently used to determine possible occupant trajectories as explained in the next section.

5.3. Trajectory determination

Occupants are often in motion, thus determining their trajectories has the potential to provide valuable information related to occupied regions on the time history. This information helps optimize important functionalities such as energy management, and security enhancement.

Trajectory determination involves the assumption that occupants walk until reaching their destinations without backtracking. The determination of the possible trajectories starts with considering all potential departure/arrival points (see Fig. 4) as possible departure spots based on the resulting CLS from the falsification process on the first captured footstep event (see Section 5.1). As explained in Section 5, when a departure point contains at least one CL within a radius of twice the distance between two footsteps, all possible trajectories corresponding to this candidate departure are taken into account. Hence, CLs of the first detected event provides all possible paths that occupants can take.

Additional detected events provide further information on paths taken from the candidate departures. CLSs related to possible departures of each new detected event are subjected to the sequential analysis as explained in Section 5.2. Resulting CLSs are then investigated to determine the possible trajectories of occupants.

In Fig. 8, a graphical representation of the trajectory-determination operation is presented. In Fig. 8, crosses represent possible departure/arrival points. CLs resulting from a sequential analysis are represented with squares, whereas the circle is the CL that is rejected using the trajectory determination. Simple dots are falsified locations.

In Fig. 8, a CL that corresponds to a possible departure (D) is rejected when its distance with at all possible arrival points (A_a) is not reduced. This operation can also be expressed mathematically by the following equation.

$$\left\{\|CL_{d,s} - A_a\|_2 - \|D - A_a\|_2\right\}_{a\in\{1,..,n_a\}} \geq 0, \forall d \in \{1, .., n_d\} \tag{5}$$

In Eq. (5), a CL that corresponds to a possible departure (D) is denoted as $CL_{d,s}$.n_a denotes the number of possible arrival points. This analysis is performed sequentially, as each CLS needs to contain locations for which the distance to at least one arrival point is reduced with respect to CLS of the previous footstep event. When CLs that correspond to a candidate departure are falsified, related trajectories are rejected. Once CLSs of all detected event signals are investigated, paths connecting remaining candidate departures with possible

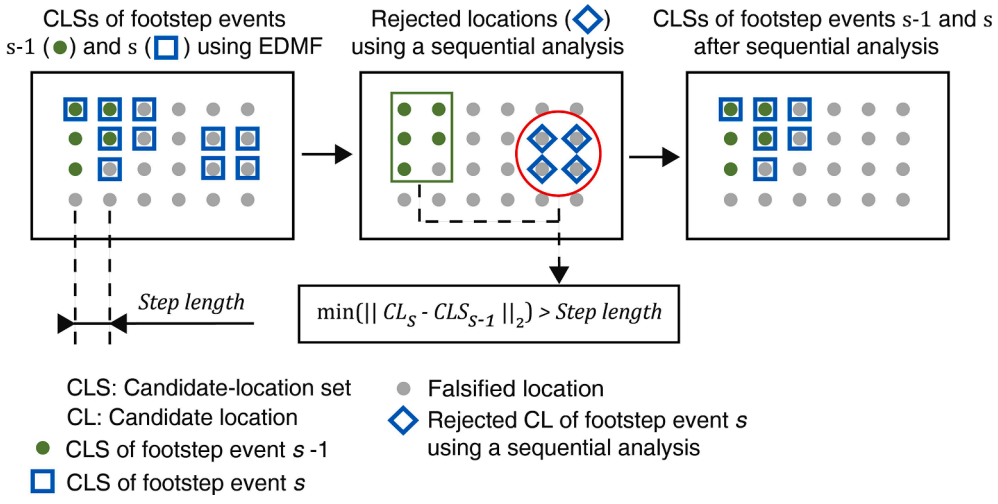

Fig. 7. Sequential analysis is applied for each CLS of each footstep event resulting from EDMF. Based on information from the previously detected footstep event, sequential analysis reduces the population of the resulting CLs.

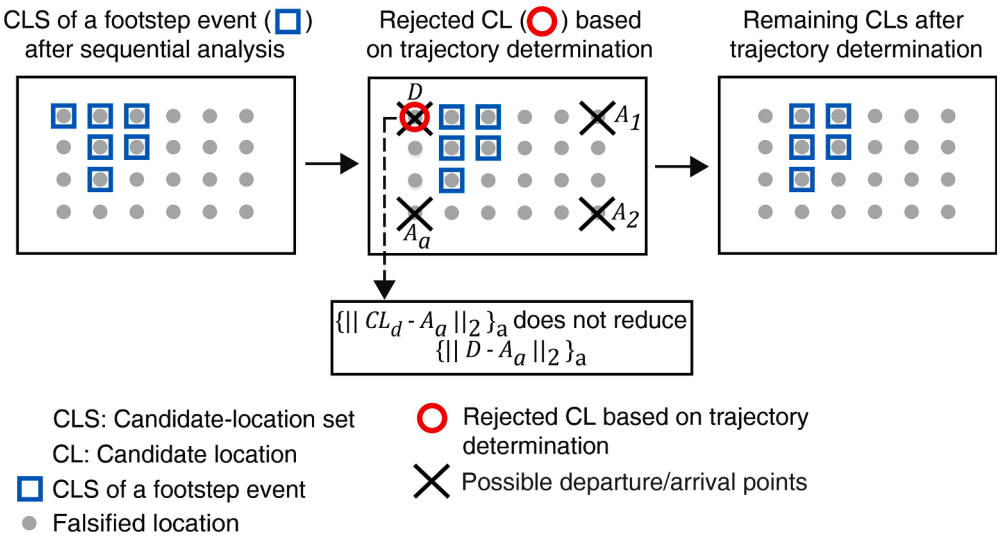

Fig. 8. A CL that corresponds to a possible departure (D) is falsified when its distances with all possible arrivals (A_a) do not reduce the distance of the corresponding departure with all possible arrivals.

arrivals that define CLs are taken to be candidate trajectories. This also ensures path continuity.

6. Case-Study 1

6.1. Description

The full-scale case study to evaluate the occupant detection and tracking framework is a floor slab of a seminar room in a multi-story building. The slab is a part of an office environment at university laboratory. The floor slab covers a surface of approximately

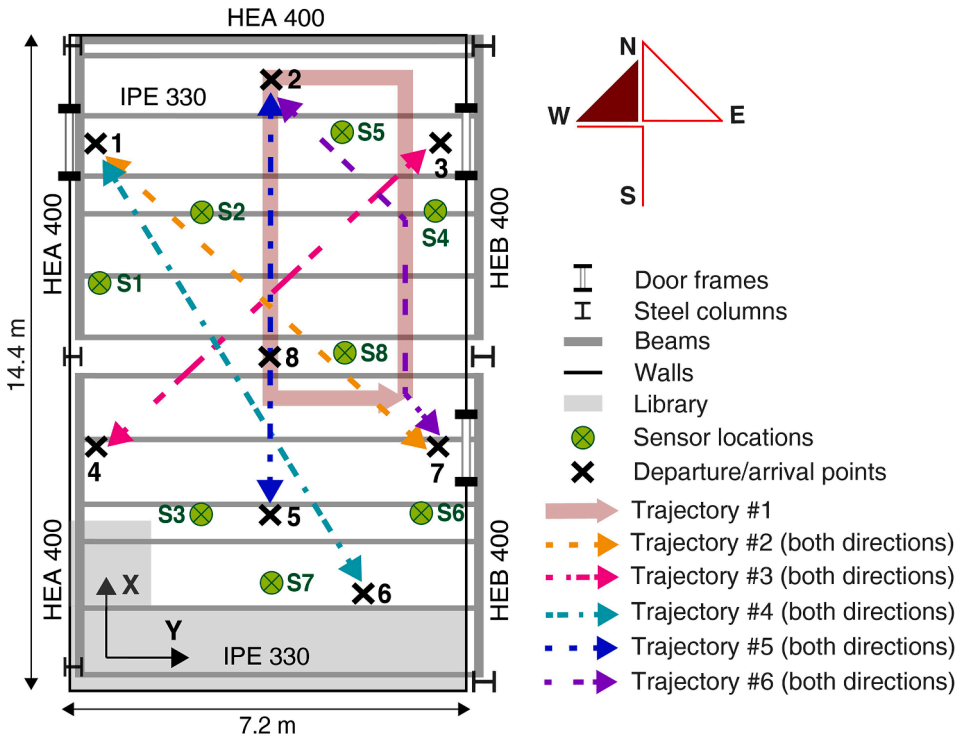

Fig. 9. Bi-directional trajectories of single occupants walking along five trajectories are used for testing. The same trajectories are used for two, three, four and five occupants walking simultaneously.

100 m^2.

The floor is composed of a reinforced concrete slab supported by steel beams, as shown in Fig. 9. The reinforced concrete slab is 20 cm thick covered by a linoleum finishing. The steel frame includes five H-beams (HEA 400 and HEB 400) and 12 I-beams (IPE 330). The floor is supported by six steel columns. The east end and upper half of the west end of the slab are connected to the remaining part of the structure (continuity of the floor slab). A non-structural wall made of plasterboard is above the structure on the east end. The lower half of the west end and the south end of the slab are connected to prefabricated reinforced concrete structural walls. The remaining parts (upper half of the west end and the north end) of the slab are joined to structural masonry walls.

The floor slab of Case Study 1 was instrumented with eight uni-directional velocity sensors (SM-24) and an acquisition unit (NI USB-6003), which is used to capture vibrations at a sampling rate of 1000 Hz. Sensor locations are chosen to cover approximatively the entire space of the floor slab (a sensor per ~ 12 m^2). Sensors S1, S4 and S6, in Fig. 9, are placed close to separation walls to increase the understanding of the support conditions. The remaining sensor positions (S2, S3, S5, S7 and S8 in Fig. 9) are placed where the dynamic responses of the first few vertical bending modes of the structure are dominant.

Walking tests are carried out for five people (O1 to O5 in Table 1) walking individually along a Trajectory #1 (see Fig. 9) to test the event detection and classification strategies, as described in Section 4. These measurements are taken for occupants walking along the trajectory without fixing the precise impact locations, without fixing the number of steps and without fixing the walking speed. All measurements are repeated several times and recorded by sensors S2, S5 and S8. Information related to occupant weights are listed in Table 1. Details of the number of walking tests for each occupant (O1 to O5 in Table 1) are given in Table 2. Using the same sensor configuration (S2, S5 and S8), vibrations are recorded for other activities, such as book-dropping, chair-dragging, hand and mug impacts on a table as well as opening and closing of doors.

Additional walking tests are recorded for occupants O1 and O5 to O8 (see Table 1) walking individually on the floor slab and following trajectories #2 to #6 (both directions; back and forth) as illustrated in Fig. 9. These new measurements are recorded on another day compared with the previous ones. All measurements are repeated several times and recorded by all sensors (see Fig. 9). Details of the number of walking tests for each occupant (O1 and O5 to O8 in Table 1) are presented in Table 2. All occupants walk while wearing various types of shoes (with either hard, intermediate or soft soles).

Measurements are recorded for two, three, four and five occupants walking simultaneously following five trajectory configurations (both directions; back and forth), as explained in Table 2. These trajectories are repeated ten times. During these walks, the occupant moves with self-selected step lengths and speeds. The walking speed (in terms of steps per second) is estimated using measurements to be between 1.5 Hz and 1.8 Hz.

Tests of single and multiple occupants following all trajectory configurations (both directions, as shown in Table 2) recorded by Sensor S2, S5 and S8 are used to test the event classification strategy described in Section 4.2. In addition, vibrations recorded by all sensors are used to test the occupant-counting strategy described in Section 4.3.

Tests that are conducted by Occupant O1 walking individually along trajectory configurations #6 to #10 in Table 2) are used for the application of model-based occupant tracking strategy described in Section 5. During these walks, Occupant O1 wears hard-soled shoes and walks at a moderate speed of approximately 1.6 Hz.

Prior modal analysis using ambient vibration measurements have revealed that the modes of the floor slab with most energy contribution to vertical bending have frequencies between 15 and 40 Hz [55]. Estimates of the fundamental bending modes of the structure are at frequencies 15.5 Hz and 24 Hz.

6.2. *Event detection*

For event detection, low-frequency components of measurements are incorporated to determine the occurrence time of possible events as explained in Section 4.1. The frequency band, 15–40 Hz, that contains the modes with most energy contribution to vertical bending of the floor slab (see Section 6.1) is divided into equivalent frequency ranges of 10 Hz with an overlap of 5 Hz. Thus, vibration measurements from the floor of the seminar room are decomposed into frequency ranges of 15–25 Hz, 20–30 Hz, 25–35 Hz and 30–40 Hz. These decompositions help focus on the frequency components that are influenced by impact events.

Event detection is successfully tested for its ability to detect more than 24,000 footstep and non-footstep events on the floor of the seminar room. For example, undetected events and incorrectly detected events are less than 1 % of the number of footstep events (1854 footstep impacts), as presented in Table 3. Footstep-event signals are extracted from the uncontrolled measurement following trajectory configurations #1 to #5 (see Fig. 9).

Table 1
Occupant weights for measurements.

Occupant	Weight (Kg)
O1	92
O2	70
O3	87
O4	67
O5	58
O6	75
O7	70
O8	72

Table 2
Trajectory combinations for one (grey fields) and multiple occupants walking simultaneously on the floor of the seminar room. O is occupant, T is trajectory and X is arrival and departure points (see Fig. 9).

Configuration	Occupant/Trajectory	Number of tests
1	**O1**: **T1** from **X2** to **X2**	19
2	**O2**: **T1** from **X2** to **X2**	13
3	**O3**: **T1** from **X2** to **X2**	16
4	**O4**: **T1** from **X2** to **X2**	9
5	**O5**: **T1** from **X2** to **X2**	9
6	**O1**: **T2** from **X1** to **X7**	6
7	**O1**: **T3** from **X3** to **X4**	6
8	**O1**: **T4** from **X6** to **X1**	6
9	**O1**: **T5** from **X2** to **X5**	6
10	**O1**: **T6** from **X7** to **X2**	6
11	**O7**: **T5** from **X2** to **X5**	10
12	**O8**: **T5** from **X2** to **X5**	10
13	**O8**: **T5** from **X2** to **X5**	10
14	**O5**: **T5** from **X2** to **X5**	10
15	**O6**: **T2** from **X1** – **O7**: **T5** from **X2**	10
16	**O6**: **T3** from **X3** – **O1**: **T6** from **X7**	10
17	**O6**: **T2** from **X1** – **O5**: **T3** from **X3** – **O1**: **T4** from **X6**	10
18	**O6**: **T2** from **X1** – **O5**: **T3** from **X3** – **O1**: **T4** from **X6** – **O7**: **T5** from **X2**	10
19	**O6**: **T2** from **X1** – **O5**: **T3** from **X3** – **O1**: **T4** from **X6** – **O7**: **T5** from **X2** – **O8**: **T6** from **X7**	10

Incorrect detection is due either to the presence of additional local maxima in $STD_{max,f}$ values or incorrect signal extraction. These additional local maxima in $STD_{max,f}$ values result in extra vibrations in the dynamic response of event signals. Extra vibrations from external sources result in incorrect determination of the ending time of detected event signal (see Section 4.1). Since each trajectory is composed of several footsteps, this detection accuracy of 99% is sufficient. By combining information from decomposed signals at several frequency ranges, the event detection strategy successfully reduces false negatives (undetected events) compared with using one frequency range, leading to accurate event detection.

6.3. Event classification

Following event detection, the next step involves differentiating between footstep and non-footstep events using the extracted signals (see Fig. 4). A binary-SVM (see Section 4.2) learning approach is used to differentiate footstep events from spurious (non-footstep) events. Binary-SVM classifier performance is compared with *k*-nearest neighbors (KNN) [46,78,79] and boosted tree (BT) [80,81] classifiers.

Good feature selection is important in order to ensure high classification performance. The use of appropriate features from event signals allows the learning algorithm to capture important patterns in the input data and to construct an efficient prediction model (see Section 4.2). Apart from frequency-domain metrics (FSV_{max} and C_{CPSD}), time-domain metrics are calculated for sensors S2, S5 and S8 (see Fig. 9) at various frequency ranges. All event signals are decomposed using CWT and reconstructed using IWT at equivalent frequency intervals of 20 Hz with an overlap of 10 Hz (see Section 4.2). This covers the frequency band, 10–240 Hz, of vibration sensors used to instrument the floor slab (see **Section 6.1**).

Time-domain metrics at specific frequency ranges that maximize the discrepancy between footstep and non-footstep event classes are selected as features for classification using null-hypothesis tests, as explained in **Section 4.2**. Based on independent footstep and non-footstep event signals, the features that are found to be useful in separating footsteps from other events are presented in Table 4.

The data set is composed of 2957 footstep events from single and multiple occupants walking simultaneously and 430 non-footstep events. Non-footstep events include book-dropping, chair-dragging, hand and mug impacts on a table, opening and closing of doors as well as jumps. The data set is randomly split into 75 % for training and 25 % for validation. This defines a supplementary validation regarding the event classification for the same case study presented by Drira et al [47].

The performance scores including accuracy, precision, recall and $F1$ of the SVM classifier, which is compared with KNN and BT classifiers are presented in Table 5. A Gaussian kernel is used to train the binary-SVM classifier since it provides better performance compared with other kernels. Classification performance of each model that is trained using raw and decomposed signals at selected

Table 3
Number of undetected and incorrectly detected events for the tested trajectory of single occupants walking on the floor of the seminar room along Trajectory #1 (see Fig. 9).

Occupant	Average number of events per test	Total number of events	Undetected events	Incorrect detection
O1	28	526	0	0
O2	28	379	1	1
O3	28	449	0	2
O4	28	242	0	1
O5	28	258	1	6

Table 4
Metrics that maximize the discrepancy between footstep and non-footstep event classes are selected as features for classification. Null-hypothesis tests of each time-domain metric (assessed at various frequency ranges) have been conducted to select the frequency ranges that best differentiates footsteps from other events.

Feature	Frequency range
σ(mean)	10–30 Hz
σ(mean)	70–90 Hz
K_r(maximum)	50–70 Hz
K_r(mean)	90–110 Hz
C_{CPSD}	All frequencies
FSV_{max}	All frequencies

frequency ranges using the null-hypothesis tests is also included for comparison.

The accuracy is defined as the number of correct predictions divided by the total number of predictions. The accuracy score is calculated using **Eq.** (6) where *TP* is number of true positives (the truth is positive, and the test predicts a positive), *TN* is the number of true negatives (the truth is negative, and the test predicts a negative), *FP* is the number of false positives (the truth is positive, while the test predicts a negative) and *FN* is the number of false negatives (the truth is negative, while the test predicts a positive).

The precision is the proportion of the correct classifications (*TP*) from all predicted positive cases ($TP + FP$). The precision score is calculated using **Eq.** (7). Recall is the proportion of the correct classifications (TP) from the number of positive cases ($TP + FN$). The recall score is calculated using **Eq.** (8). The *F*1score is the overall performance metric that reflects the classifier model's ability to distinguish between classes. The *F*1 score is calculated using **Eq.** (9).

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = 2\frac{Precision^{*}Recall}{Precision + Recall} \tag{9}$$

Type I and II errors are used for further comparison of classification approaches. Type I error is the rate of footstep events that are identified as non-footstep events. Type II error is the rate of non-footstep events that are identified as footstep events.

In Table 5, the binary-SVM classifier, trained using raw footstep-event signals from single and multiple occupants walking simultaneously, provides similar overall performances compared with KNN and BT classifiers. For example, the overall performance as defined by *F*1 score (see **Eq.** (9)) exceeds 97 % in most cases. However, all classifiers have type II errors (more than 19 %). Thus, using raw signals for training leads to non-reliable classifiers that are unable to distinguish spurious events from footstep events.

Binary-SVM, KNN and BT classifiers, trained using decomposed event signals (from single and multiple occupants walking

Table 5
Classification performance, based on validation test (25 % of data set), to distinguish between footstep events (single and two to five occupants walking simultaneously) and non-footstep events. Performance of classifiers that are trained using raw and decomposed signals at selected frequency ranges using the null-hypothesis tests is included for comparison.

		SVM	KNN	BT
Raw signals	Accuracy (%)	95.2	95.7	96.9
	Precision (%)	95.5	96.2	97.2
	Recall (%)	99.1	99.1	99.3
	*F*1(%)	97.3	97.6	98.3
	Error I (%)	1.0	1.0	0.7
	Error II (%)	30.9	26.4	19.1
Decomposed signals (CWT)	Accuracy (%)	98.8	98.2	98.2
	Precision (%)	99.1	98.5	99.1
	Recall (%)	99.6	99.5	98.9
	*F*1(%)	99.3	99	99
	Error I (%)	0.4	0.5	1.1
	Error II (%)	6.4	10.1	6.4

simultaneously) using CWT and IWT at selected frequency ranges, provide a marginal improvement of prediction performances compared with those assessed with raw signals. All performance scores, including accuracy precision, recall and *F*1 score (see **Eq. (6) to (9)**), exceed 98 %.

Moreover, as shown in Table 5, feature selection using null-hypothesis tests at various frequency ranges have led to significant reduction of type II error. Therefore, event-classification performance is enhanced by selecting appropriate frequency components of vibration measurements. Event classification using binary-SVM is less sensitive to type I and type II errors compared with KNN and BT classifiers. Thus, the binary-SVM classifier provides better event classification than KNN and BT classifiers.

6.4. Occupant counting

Differentiating between the number of occupants walking on the floor of the seminar room is achieved using a SVM classifier, as explained in **Section 4.3**. Cross-correlation coefficients between signals measured at various sensor locations from a same footstep event are used as features to train the SVM classifier. Apart from cross-correlation coefficients, standard deviation (σ) values of event signals recorded at each sensor location and maximum CPSD of all sensors are used as features to enhance the performance of the occupant-counting strategy [47].

Occupant counting has been tested with footstep-event signals from five participants (O1 and O5 to O8 in Table 1) that have been walking individually and together including two, three, four and five occupants and following multiple trajectories (see Table 2). Occupant-counting classification is performed using three classification strategies to differentiate between: 1) one and more than one occupant, 2) one and two occupants and 3) one, two, three, four and five occupants. When the first and the second classification strategies are used to count the number of occupants, they are developed using a binary-SVM classifier on raw footstep-event signals. This study provides a supplementary validation for results from another case study presented previously by Drira et al [47]. The third classification strategy is developed using a multi-class SVM classifier. SVM classifiers to determine the number of occupants are trained based on several kernels, including linear kernel, Gaussian kernel, third-and-fourth degree polynomial kernels. The third-degree polynomial kernel provides best classification performance compared with other kernels.

Binary-SVM classifier performance is compared with KNN and BT classifiers, in Table 6. The data set for the first model includes 1129 footstep events from single occupants and 695 events from more than one occupant (two to five occupants) walking together. The data set for the second model includes 1129 footstep events from single occupants and 256 events from two occupants walking together. The data set for each classifier is randomly split into 75 % for training and 25 % for validation.

Accuracy, precision, recall and F1 scores, calculated using **Eq. (6)** to **Eq. (9)** are presented in Table 6. These metrics help to assess the performance scores for each classifier using raw footstep-event signals. Also, type I error that defines the rate of one occupant classified as more than one occupant and type II error that defines the rate of more than one occupant classified as one occupant are used for further comparison between classification approaches.

Based on validation test (25 % of data set), the binary-SVM classifiers are able to differentiate between the presence of either one, more than one occupant or two occupants with performance scores exceeding 93 % as shown in Table 6. Binary-SVM classifiers provide better performance scores (average increase of 5%) than KNN and BT classifiers. For example, the overall prediction performances defined by F1 score of the first and the second models are equal approximately to 95 % and 97 % for SVM, 87 % and 93 % for KNN and 89 % and 94 % for BT classifier (see Table 6).

Significantly fewer type II errors (two occupants or more than one occupant are classified as one occupant) are produced using SVM (13 % and 17 %) compared with KNN (33% and 42%) and BT classifiers (29% and 36%). Type I error is also lower when using SVM classifier compared with KNN and BT. Using cross-correlations between footstep signals, standard deviation (σ) values of event signals recorded at various sensors and maximum CPSD of all sensors as features provide important classification performance to distinguish between the presence of either one or more than one occupant, also either one or two occupants.

Performance of the multi-class SVM classifier (classification strategy 3) to determine the number of occupants (up to five) are also compared with KNN and BT classifiers as illustrated in Table 7. The data set for the third model includes 1129 footstep events from single occupants, 256 events from two occupants, 143 events from three occupants, 153 events from four occupants and 143 events

Table 6
Classification performance, based on validation test (25 % of data set), to distinguish between two classification strategies: 1) one and more than one occupant and 2) one and two occupants on the floor of the seminar room (see Fig. 9).

		SVM	KNN	BT
1) One or more than one occupant	Accuracy (%)	93	82	84.9
	Precision (%)	93.1	83.6	85.4
	Recall (%)	96.3	89.9	92.6
	F1 (%)	94.7	86.6	88.8
	Error I (%)	3.7	10.1	7.4
	Error II (%)	13.1	32.5	29.4
2) One or two occupants	Accuracy (%)	94.2	87.6	90.2
	Precision (%)	96.1	90.8	92.2
	Recall (%)	96.8	94.3	96.1
	F1 (%)	96.5	92.5	94.1
	Error I (%)	3.2	5.7	3.9
	Error II (%)	17.2	42.2	35.9

from five occupants walking together. The data set for each classifier is randomly split into 75 % for training and 25 % for validation.

Using validation testing (25 % of data set), the multi-class SVM classifier is able to differentiate between the number of occupants (up to five occupants) with an average accuracy of approximately 84 % as shown in Table 7. The multi-class SVM classifier provides better average performance scores (average increase between 12 % and 39 %) than KNN and BT classifiers.

The normalized confusion matrix of five-occupant counting resulting from the multi-class SVM classifier is shown in Table 8. The normalized values define the proportion of correct classifications of the number of occupants from all participants (i.e. recall score in **Eq. (8)**).

It can be observed in Table 8, that the multi-class SVM classifier is able to differentiate between one and more than one occupant with a high recall score of 96.5 %. Similar results are found for the binary-SVM classier as shown in Table 6. The multi-class SVM classifier is also able to determine the number of five occupants walking on the floor with a significant recall score of 80.6 %. However, the counting classifier results in low-performance scores for two, three and four occupants. For example, the determination of two occupants walking together on the floor is confused with one occupant ~ 23 % of the time. The number of three occupants on the floor is confused with one and two occupants, ~20 % and ~ 23 % of the times, whereas, the number of four occupants is confused with five occupants ~ 26 % of the time.

To conclude, using cross-correlation coefficients between event signals at all sensor locations, standard deviation (σ) values of event signals at each sensor location and maximum CPSD of all sensors as features improves the performance of the classifier for distinguishing between the presence of either one or up to five occupants on the floor. The occupant-counting strategy is successfully validated for two full-scale floor slabs.

6.5. Occupant tracking

Tracking of a single occupant is achieved using the model-based approach that is described in **Section 5**. Eight departure/arrival points are fixed for this case study, denoted by crosses (X1 to X8) in Fig. 9. These points lead to 56 possible trajectories. Vibration measurements are conducted by Occupant O1 walking along trajectories #2 to #6 (both directions; back and forth) as presented in Table 2 (Configurations #6 to #10). Walks along these trajectories are repeated six times. Footstep-impact simulations are generated using a finite element model of the floor slab, as described in **Section 6.5.1.** The simulation results are used in the application of model-based occupant localization using EDMF (see **Section 5.1**).

6.5.1. Model predictions

For the application of occupant localization using a model-based approach, as explained in **Sections 5**, physics-based models are incorporated in the interpretation of vibration measurements to identify occupant locations. Footstep-impacts are simulated using a finite element of the floor slab of the case study. Linear modal superposition is used to calculate the dynamic response caused by footstep impacts using ANSYS [82].

The finite element model of the floor slab is modeled using shell elements (SHELL181) and beams are modeled as beam elements (BEAM188). Beam elements are assumed to be fully connected to the shell elements. Also, columns are modeled as simple supports (see Fig. 9). The elastic moduli for the steel and the concrete slab are taken to be 210 and 35 GPa.

Due to incomplete knowledge of boundary conditions of the floor slab, the separation walls (see Fig. 9) are modeled using translational zeros-length springs in the vertical direction (COMBIN14). Four spring elements are used in the finite element model to describe the upper half of the west end of the slab (masonry wall), the lower half of the west end and the south end of the slab (reinforced concrete walls), the east end of the slab (plasterboard walls), and the north end the slab (masonry wall that is connected to a concrete staircase), see Fig. 9.

Stiffness values of these springs are estimated based on a prior sensitivity analysis. The Latin-hypercube sampling approach [83] is used to generate 500 spring-stiffness values from sufficiently small to sufficiently large. Values of each spring element are varied at a time using modal analysis simulations. This analysis results in an s-shaped function of the fundamental frequency as a function of each spring stiffness. The stiffness values between freely supported and completely fixed of all spring elements are 316 N/mm, 631 N/mm, 1259 N/mm and 200 N/mm respectively.

A load model, in the time domain as illustrated in Fig. 10, that approximates a typical vertical ground-reaction force (VGRF) due to a footstep is applied to the finite element model at possible locations [55,70,75]. After excluding the occupied space by the small library in the room (see Fig. 9), 90 % of the floor slab is divided into a grid of possible locations. The distance between two possible locations is assumed to be 37.5 cm (half of the assumed step length), leading to 612 possible footstep locations.

The load model, in Fig. 10, defines the three phases of the VGRF: 1) the heel phase, 2) the heel-to-toe phase and 3) the toe-off phase

Table 7
Average classification performance of the third classification strategy, based on validation test (25 % of data set), to determine the number of occupants (up to five occupants) on the floor of the seminar room (see Fig. 9).

	SVM	KNN	BT
Accuracy (%)	83.8	74.8	72.6
Precision (%)	72.3	62.5	57.8
Recall (%)	70.2	56.1	50.6
F1 (%)	71.0	58.4	51.8

Table 8
Normalized confusion matrices of five-occupant counting. Recall results are based on validation test (25 % of data set). Occupant-counting classification has been performed using a multi-class SVM classifier.

	Recall (%)					
True label	1	96.5	1.4	1.4	0.7	0
	2	23.4	60.9	14.1	0	1.6
	3	20	22.9	51.4	2.9	2.9
	4	2.6	5.1	5.1	61.5	25.6
	5	0	0	0	19.4	80.6
		1	2	3	4	5
	Predicted label					

as explained in detail by Racic et al. [68] and Drira et al [55,70,75]. From Fig. 10, the heel phase starts with an initial heel-contact (IHC) phase that is characterized by a brief duration, denoted as T1. During this phase, an abrupt transfer of the part of the body-weight, denoted as F1, to the ground is achieved. The heel phase ends with a full heel-contact (FHC) phase during which the foot is in full contact with the ground until the VGRF reaches a maximum, denoted as F3. F2 refers to the attenuation in magnitudes within the heel phase. T2 refers to the duration of this phase.

The heel phase is followed by the heel-to-toe phase, during which the opposite foot leaves the ground for the next footstep impact whereas the heel of the stance foot starts to rise from the floor surface. This explains the descending trend that defines a minimum magnitude denoted as F4. T3 refers to the duration of the heel-to-toe phase. Once achieved, the foot contact is completely transferred to toes, and the opposite foot heel is in contact with the floor surface. This explains an ascending trend to a maximum magnitude denoted as F5. Finally, the toe-off phase refers to the rising of the stance foot during which VGRF presents a decreasing trend to zero. T4 refers to the duration of the toe-off phase.

The footstep-load function (f_{load}) that is proposed by Drira et al [55,70,75] is defined by **Eq. (10)**. This function is constructed using four sine functions and a cosine function to represent the VGRF presented in Fig. 10. In **Eq. (10)**, forces F1 to F4 and durations T1 to T3 define the load function parameters. T presents the total duration of the VGRF.

$$f_{load}=\begin{cases} F1\,sin\left(\frac{2\pi t}{T1}\right) & if \quad t\in[0\ldots T1]\ s \\ F2+\frac{F1-F2}{2}\left(1+sin\left(-\frac{4\pi}{T2}\left(t-T1+\frac{3T2}{8}\right)\right)\right) & if \quad t\in\left[T1\ldots T1+\frac{T2}{4}\right]s \\ F2+\frac{F3-F2}{2}\left(1+sin\left(\frac{4\pi}{3T2}\left(t-T1+\frac{7T2}{8}\right)\right)\right) & if \quad t\in\left[T1+\frac{T2}{4}\ldots T1+T2\right]s \\ F4+\frac{F3-F4}{2}\left(1+cos\left(\frac{2\pi}{T3}\ (t-T1-T2)\right)\right)+(t-T1-T2)\ \frac{F5-F3}{T3} & if \quad t\in[T1+T2\ldots T-T4\ ]\ s \\ \frac{F5}{2}\left(1+sin\left(-\frac{\pi}{T3}\left(t-\frac{5}{2}(\ T-T4)\right)\right)\right) & if \quad t\in[\ T-T4\ldots T\ ]\ s \end{cases} \tag{10}$$

Average parameter values defining forces F1 to F5 and durations T and T1 to T3 that are used to apply the load model (see **Eq. (10)**) for simulations are presented in Table 9. These values are determined based on prior analysis of measured VGRFs from multiple

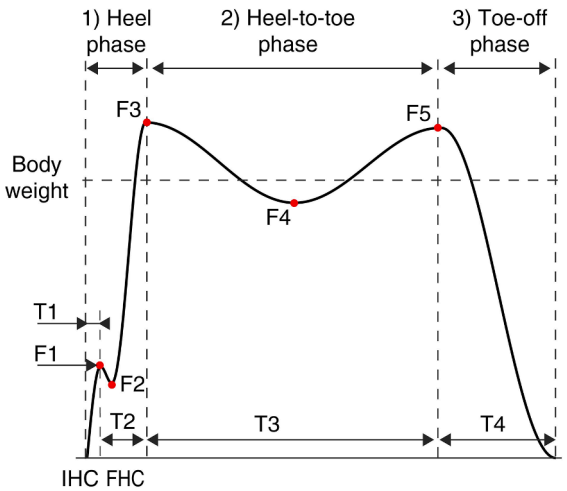

Fig. 10. Vertical footstep-impact load model.

occupants [55,70,75].

6.5.2. Uncertainty estimation

Model simulations are prone to uncertainties from sources such as model imperfections (idealized boundary conditions and omissions), unknown model parameters and idealized footstep-impact load functions. The finite element model of the floor-slab (see **Section 6.5.1**) involves many simplifications, including the use of shell elements for the concrete slab and one-dimensional bar elements for the supporting beams. Also, model simplification includes the use of translational springs (rotation free) to model the plasterboard, masonry and reinforced concrete walls (see Fig. 9). The finite element model does not include elements such as the room furniture, linoleum floor finishing and connections (in the horizontal direction) between the floor slab and the reinforced concrete walls (see **Section 6.5.1**).

In addition, the footstep-impact load function [70,75] is applied independently to a single node in simulations at each predefined location. This excludes the possible contribution of the other foot during the pre-swing phase of the gait pattern (when two feet are in contact with the ground) [70,75]. This further increases the uncertainty associated with the simulation model. Based on engineering judgement and previous work [36,84,85], the uncertainty related to model simplifications and omissions is estimated to be uniformly distributed between − 15% to + 25% of simulated velocity amplitudes.

According to a sensitivity analysis of the applied footstep-load function [70,75], the heel phase is the most important stage. The frequency of the heel phase (1/duration) induces low-frequency response components [36]. The simulation is performed for responses near the natural frequencies of the structure. Underestimation of friction effects results in over-estimated simulated velocity amplitudes, thereby leading to additional model uncertainties of −30 % to 0 %.

Measurements are affected by uncertainties originating from sensor resolution, precision and variations in floor vibrations due to natural variability in walking gaits (see **Section 2**). Sensor resolution and precision provided by the sensor manufacturer are not significant (approximately 2 %). Inherent variability in walking gait, resulting from several individuals defining various anatomies and walking styles, walking at various speeds and wearing various type of shoes are determined prior to model-based occupant localization [55,70]. For this case study variability in walking gait is approximately bounded between −72 % and + 54 % [55,70] which defines the 99th percentile of the uncertainty distribution.

Subsequently, model and measurement uncertainties related to each detected footstep event are combined using Monte-Carlo sampling with one million samples (see **Section 5.1**). Using a target reliability of localization of 95%, localization thresholds for each detected footstep event are derived from the combined uncertainty distribution according to **Eq. (3)**.

6.5.3. Occupant trajectories

Prior to the application of model-based occupant-localization, the measured and simulated footstep-impact signals are decomposed using CWT and reconstructed using IWT at a frequency range of 15–40 Hz (see **Section 6.1**). This range contains the modes of the floor slab with most energy contribution to vertical bending. The standard deviations (σ) of the filtered signals are used as a metric for the model-based occupant localization strategies.

For each detected footstep-event, occupant localization is carried out through explicitly including systematic errors and model bias using EDMF to identify a population of possible locations denoted as CLS as explained in **Section 5.1**. Taking into account information from previously detected footstep events, sequential analysis is carried out to enhance the precision of CLS of each footstep event as described in **Section 5.2**. Sequential analysis reduces the ambiguity of CLS of each footstep event by verifying that the distance between two successive footstep events does not exceed a predefined distance (pre-defined step length). CLSs resulting from the model-falsification approach and a sequential analysis for each detected footstep are then used to identify occupant trajectories. A trajectory determination is performed as described in **Section 5.3**.

An example of CLSs of a few events (first, intermediate and last events) resulting from Occupant O1 walking along each trajectory (see Fig. 9) are shown in Fig. 11. In Fig. 11, squares represent the CLSs, and dots represent the falsified location sets. Dashed lines represent the separation walls, and h-shapes represent the steel columns. Diamonds represent sensor locations, and real footstep locations of each occupant are represented with crosses.

Regarding real footstep locations in Fig. 11, incorporating physics-based models in the interpretation of event signals from a single occupant using EDMF has led to accurate localization results for all events. Moreover, a sequential analysis that accounts for

Table 9
Average parameter values for footstep-impact load function.

Load-model parameter		Parameter value
F1	Initial heel-contact force (Kg)	30
F2	Initial-to-full heel-contact force (Kg)	22.5
F3	Full heel-contact force (Kg)	93.7
F4	Heel-to-toe contact force (Kg)	70.3
F5	Toe contact force (Kg)	86.3
T1	Initial heel-contact duration (s)	0.025
T2	Full heel-contact duration (s)	0.12
T3	Heel-to-toe contact duration (s)	0.45
T	Footstep-contact duration (s)	0.8

information from previous events, and trajectory determination have proved to increase localization precision. In Fig. 11, the CLSs of the last few events are smaller compared with the first few events and this is achieved without compromising accuracy. The average localization accuracy of all measurements (30 walking tests), resulting from the model-based occupant tracking approach (see **Section 5**) is approximately 94 %. The average localization precision is approximately 71 %.

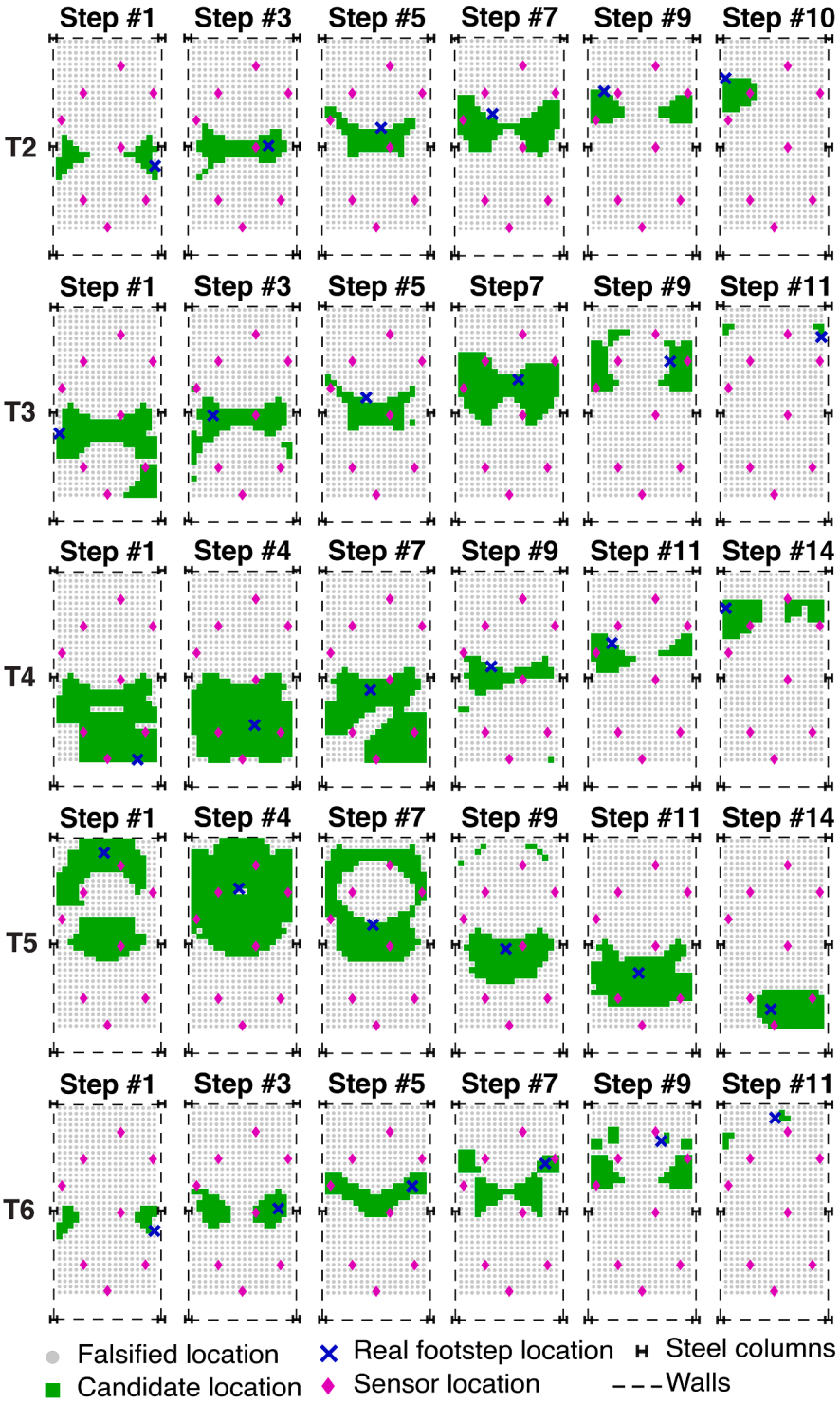

Fig. 11. CLSs that are obtained using EDMF, sequential analysis and trajectory determination on the floor of Case Study 1. For each footstep event, represented CLS corresponds to all remaining trajectories. Trajectory is denoted as T.

In comparison with the localization results obtained using a two-sensor configuration, as shown by Drira et al [75], increasing the number of sensors for the model-based localization enhances precision. Using an eight-sensor configuration to cover the entire floor space provides more information than a two sensor configuration for the falsification process using EDMF. The rate of the rejected location instances is more important than those obtained using only two sensors. For this application, the average increase of the localization precision is approximately 51 % compared with localization results (using EDMF and sequential analysis) found by Drira et al [75] for two sensors.

In Fig. 11, model-based occupant tracking has led to accurate determination of correct arrival points for all tested trajectories. For Trajectory #2 (Fig. 9), the occupant requires 10 footstep events to attain the arrival point X7 or X1, while for trajectories #2 and #5, the occupant requires 11 footstep events to attain the arrival points X2, X3, X4 or X7. For trajectories #3 and #4, the occupant requires 14 footstep events to attain the arrival points X1, X2, X5 or X6.

Assuming that all departure/arrival points (X1 to X8 in Fig. 9) are predefined, the model-based occupant-tracking operation provides accurate and precise candidate trajectories for an occupant, as presented in Table 10. Table 10 summarizes the number of candidate trajectories at the last detected footstep event of tested trajectories #2 to #6 (Fig. 9) repeated three-times, back and forth.

From Table 10, the average tracking accuracy is 90 % in which the correct paths for three among thirty walking tests are not determined once all detected events are investigated. An average tracking precision of 97 % is observed for all tested measurements. Tracking precision refers to the percentage of falsified trajectories from all possible paths (56 possible trajectories). For example, when only one possible trajectory remains unfalsified, then occupant tracking is taken to be 100 % precise.

Therefore, combining the knowledge of structural behavior with measurements and explicitly taking into account sources of uncertainties, model-based occupant tracking has the potential to provide accurate and precise candidate trajectories of an occupant walking on a full-scale floor slab.

7. Case-Study 2

7.1. Description

The objective of this case study is to evaluate the model-based tracking strategy described in **Section 5**. This evaluation involves tracking single and two occupants walking simultaneously along more realistic scenarios.

The full-scale floor slab is an open-space hall in a multi-story building, as shown in Fig. 12. The floor is a continuous reinforced-concrete slab that covers an area of approximately 950 m^2. The effective tested area is approximately 600 m^2, as shown in Fig. 12. Ten concrete columns, as well as several reinforced-concrete walls, support the floor slab. The concrete slab is 25 cm thick and is covered by a linoleum finishing. Uni-directional reinforced-concrete beams connect the slab with the concrete columns (see section A-A in

Table 10
Model-based occupant-trajectory results.

Trajectory	Path	Walking test (#)	Detected events (#)	Candidate trajectories at the last event	Correct trajectory	Average precision (%)
T2	**X1 to X7**	1	10	3 out of 56		**98.2**
		2	10	1	–	
		3	10	3		
	X7 to X1	1	10	1		
		2	10	2		
		3	10	2		
T3	**X3 to X4**	1	11	3		**96.4**
		2	11	3		
		3	11	3		
	X4 to X3	1	11	4		
		2	11	3		
		3	11	2		
T4	**X6 to X1**	1	14	2		**96.7**
		2	14	4		
		3	14	2		
	X1 to X6	1	14	3		
		2	14	3		
		3	14	3		
T5	**X2 to X5**	1	14	3		**97**
		2	13	3		
		3	14	1		
	X5 to X2	1	14	3	–	
		2	14	3		
		3	14	3		
T6	**X7 to X2**	1	11	4		**97**
		2	11	2		
		3	11	1	–	
	X2 to X7	1	11	4		
		2	11	2		
		3	11	3		

Fig. 12). Several masonry and plasterboard walls are used to separate the hall into functional spaces.

The floor is instrumented with the same vibration sensors (uni-directional Geophones SM-24 by I/O Sensor Nederland) and acquisition unit (NI USB-6003) as the previous case study (**Section 6**) to measure vertical velocity-response of the slab with a sampling rate of 1000 Hz. Eight sensors (one sensor per ~ 75 m^2) are placed at locations where dynamic responses of the first few vertical bending modes of the structure are dominant.

Measurements are recorded for three occupants walking on the floor slab individually along six trajectories (both directions; back and forth) as illustrated in Fig. 12, where crosses (X1 to X12) represent possible departure/arrival points. All occupants walk while wearing various types of shoes (with hard, intermediate or soft soles). The three occupants (O1 to O3) weigh 93 Kg for O1, 87 Kg for O2 and 75 Kg for O3. A priori knowledge of occupant characteristics (weigh, type of shoes and walking speed) during testing has not been required. Moreover, measurements are recorded for two occupants walking simultaneously following six trajectory configurations (both directions; back and forth). See configurations 7 to 12 in Table 11 for details regarding the trajectories involved in two occupants walking simultaneously on the floor.

Single occupants walking along six trajectories, as well as six trajectory configurations of two occupants walking together (see Table 11) are used for testing the strategies described in **Section 5**. Walks along these trajectories are repeated several times, as shown in Table 11. During these walks, the occupant moves with self-selected step length and speed in order to provide realistic walking scenarios. The walking speed (in terms of steps per second) for individual occupants is estimated to be between 1.5 Hz and 1.8 Hz using measurements. For all measurement tests involving two occupants, the walking speed is approximately 1.6 Hz.

Based on ambient vibration measurements, the modes of the structure with most energy contribution to vertical bending have frequencies between 5 and 30 Hz. The fundamental bending mode of the structure is contained within the frequency range of 9–11 Hz.

7.2. Occupant tracking

The model-based tracking strategy is intended to identify possible locations of occupants walking on the floor slab using error-domain model falsification (EDMF), as explained in **Section 5.1**. These locations are then used to determine possible trajectories of occupants. Using EDMF, location instances that do not contradict measured vibrations at all sensors define the candidate-location set (CLS). This operation is repeated separately for each detected event signal induced by occupant footsteps.

Simulated and measured vibrations are filtered using CWT at frequency range of 5–30 Hz in (see **Section 7.1**) to increase their SNR. Maximum difference in amplitudes (Δ_{amp}) and standard deviation (s) of processed event signals are used as metrics for the model-based occupant localization since they are well-suited to evaluate the localization process [38].

For this application, floor-vibration measurements from a single and two occupants walking simultaneously as presented in Table 11 are investigated. In the context of the model-based tracking strategy, twelve departure/arrival points are fixed in this case

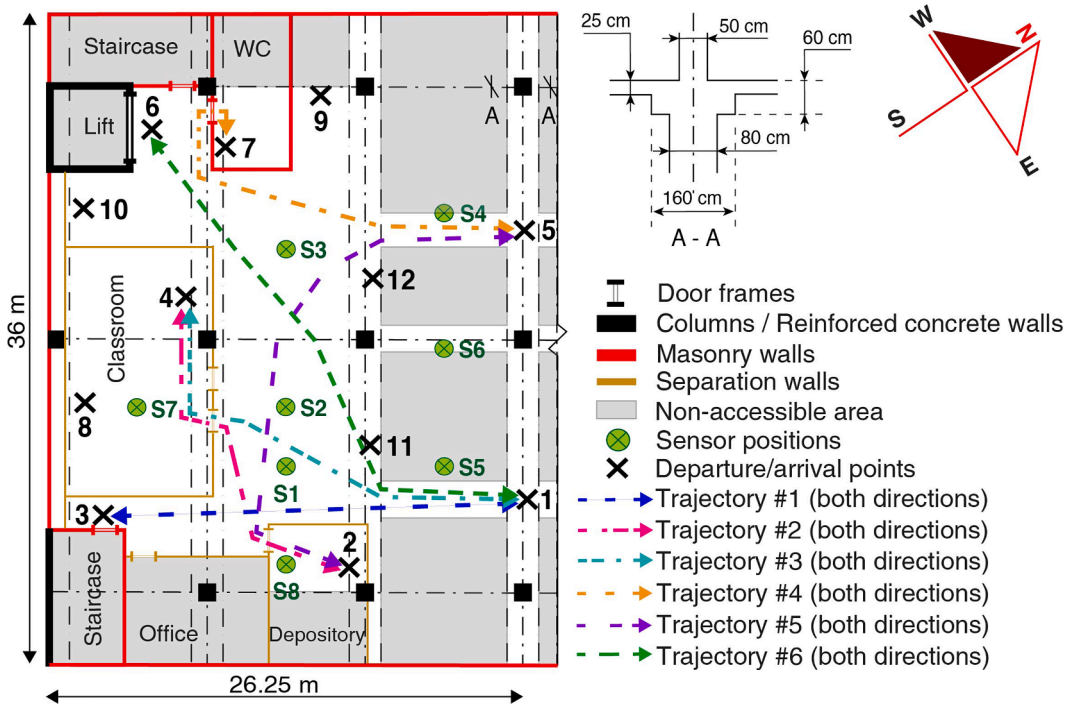

Fig. 12. Model-based occupant tracking is tested on a full-scale concrete slab (~950 m2). Bi-directional trajectories of single occupants walking along six trajectories are used for testing. The same trajectories are used for two occupants walking simultaneously.

Table 11
Trajectory configurations for one (grey fields) and two occupants walking simultaneously on the floor of the open-space hall in Singapore. For each configuration, walking tests are repeated for both directions (back and forth). O is occupant, T is trajectory and X is arrival and departure points (see Fig. 12).

Configuration	Occupant/Trajectory	Number of tests
1	**O1**: **T1** from **X1** to **X3**	8
	O2: **T1** from **X1** to **X3**	6
	O3: **T1** from **X1** to **X3**	6
2	**O1**: **T2** from **X2** to **X4**	10
3	**O1**: **T3** from **X1** to **X4**	8
	O2: **T3** from **X1** to **X4**	4
4	**O1**: **T4** from **X5** to **X7**	12
	O3: **T4** from **X5** to **X7**	4
5	**O1**: **T5** from **X2** to **X5**	6
	O2: **T5** from **X2** to **X5**	4
6	**O1**: **T6** from **X1** to **X6**	8
	O2: **T6** from **X1** to **X6**	4
7	**O1**: **T6** from **X6** - **O3**: **T3** from **X1**	6
8	**O1**: **T2** from **X2** - **O3**: **T1** from **X1**	6
9	**O1**: **T3** from **X4** - **O3**: **T3** from **X1**	6
10	**O1**: **T5** from **X5** (After ~ 4 FSs) - **O3**: **T5** from **X5**	3
	O1: **T5** from **X5** (After ~ 6 FSs) - **O3**: **T5** from **X5**	4
	O2: **T5** from **X5** (After ~ 4 FSs) - **O3**: **T5** from **X5**	2
11	**O1**: **T3** from **X1** (After ~ 4 FSs) - **O2**: **T1** from **X3**	6
12	**O2**: **T3** from **X1** (After ~ 4 FSs) - **O1**: **T3** from **X1**	6

study, as shown by crosses (X1 to X12) in Fig. 12. These points lead to 132 possible trajectories.

CLSs of detected event signals are used to determine candidate trajectories of occupants as shown in **Section 5.1**. Candidate locations (CLs) resulting from the first detected event help determine candidate departures. CLs associated with each candidate departure point are then used to investigate possible paths based on information from succeeding detected footstep events.

The CLS related to each succeeding detected event is then subjected to a sequential analysis to reduce ambiguities in localization results. The sequential analysis, as explained in **Section 5.2**, involves the assumption that the distance between two successive impact events cannot exceed a predefined distance. As impact locations during walking tests are not defined, this distance is assumed to be twice the pre-defined step length. For each newly detected event signal, the sequential analysis is applied separately for each CLs associated with each candidate departure. Finally, CLSs of each detected event, resulting from the sequential analysis, are investigated to determine possible trajectories.

The trajectory-determination operation, as explained in **Section 5.3**, involves the assumption that occupants walk until reaching their destinations without backtracking. Hence, a CL that corresponds to a possible departure is rejected when its distance to at least one possible arrival point is not reduced. Once CLSs of all detected event signals are investigated, paths connecting remaining candidate departures with possible arrivals that define CLs are taken as candidate trajectories.

7.2.1. Model predictions for a single- and two-occupant localization

Prior modal analysis based on ambient vibration measurements has revealed that the first bending mode of the floor slab lies within the frequency range of 9–11 Hz. This allows classification of the full-scale slab as a low-frequency floor according to refs [86–88]. For low-frequency floors, a deterministic vertical load model for a single footstep impact is expressed in the time domain by the summation of harmonic components [89] (Fourier series) as defined by Blanchard et al. [90] and described in the following equation:

$$F(t) = G + \sum_{i=1}^{n} G\alpha_i \sin(2\pi i f_p t - \varphi_i) \tag{11}$$

where $F(t)$ is presented by a static part expressed as G which is the person's static weight and a fluctuating part expressed by the harmonics. f_p is the walking frequency, which is approximately between 1.4 Hz and 2.5 Hz [68,91]. t is time. φ_i is the phase shift of the i^{th} harmonic. α_i is the Fourier coefficient, also known as the dynamic load factor (DLF) of the i^{th} harmonic. n is the number of contributing harmonics.

For this application, the walking frequency is fixed at 1.6 Hz (i.e. duration of 0.625 s), and the phase shift of the i^{th} harmonic is assumed equal to zero. Also, the occupant weight is assumed equal to 85 Kg, which corresponds approximately to the average weight of the participants (see **Section 7.1**). The DLFs that are involved in the footstep-impact simulations are proposed by Young [92], as defined by **Eq. (12)**. In **Eq. (12)**, the resulting DLFs are for the first four harmonics, which are ascertained as regression equations as a function of walking frequency (f_p). The walking frequency was assumed to vary between 1 and 2.8 Hz [92].

$$\begin{aligned} \alpha_1 &= 0.41\left(f_p - 0.95\right) \le 0.56 \quad f_p = 1 - 2.8Hz \\ \alpha_2 &= 0.069 + 0.056 f_p \qquad f_p = 2 - 5.6Hz \\ \alpha_3 &= 0.033 + 0.0064 f_p \qquad f_p = 3 - 8.4Hz \\ \alpha_4 &= 0.013 + 0.0065 f_p \qquad f_p = 4 - 11.2Hz \end{aligned} \tag{12}$$

Footstep-impact simulations are generated using a finite-element model of the slab. The footstep-impact force is applied to a single node as a function of time. The dynamic responses of the slab are generated based on the linear-modal-superposition analysis in ANSYS [82]. The floor slab is modeled using solid elements (SOLID185). The elastic modulus for the concrete slab is taken to be 35 GPa. The viscous damping ratio is taken conservatively to be 5%. Columns and reinforced-concrete walls are modeled as simple supports (see

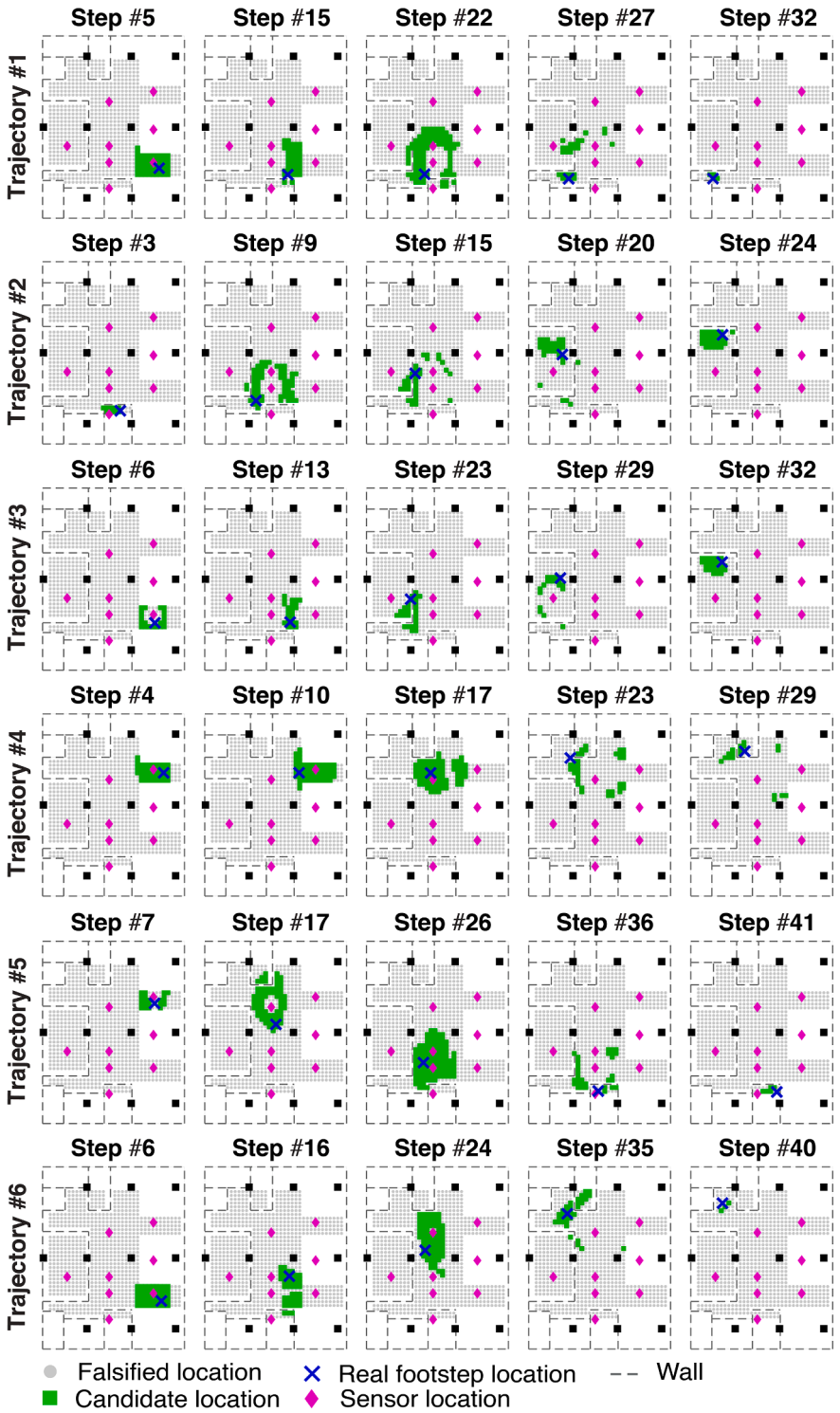

Fig. 13. CLSs that result from occupants walking individually along trajectories #1 to #6 (see Fig. 12) are obtained using EDMF, sequential analysis and trajectory determination.

Fig. 12).

Two-thirds of the floor slab is divided into a grid of possible locations (see the accessible area in Fig. 12). The distance between two possible locations is assumed to be 75 cm (see **Section 5**), leading to 796 possible footstep locations. These simulations are used for the model falsification to localize single occupants.

Also, these footstep-impact simulations are used to generate possible model predictions associated with two occupants walking simultaneously on the floor slab, as explained in **Section 5.1**. Contributions from each occupant at each possible location are superimposed based on two time-offsets (0.2 s and 0.4 s, as explained in **Section 5.1.2**). The generation of model predictions involves the assumption that two occupants cannot be at the same location at a time. Including contributions from a single occupant, simulated responses for two occupants result in 1,266,436 model instances (796x795x2 + 796).

7.2.2. Uncertainty estimation

Similar to the previous case study (see **Section 6.5**), the finite element model does not include elements, such as separation walls, room furniture, and connections between the floor slab and the reinforced concrete walls. Also, unknown model parameters and an idealized footstep-impact load function increase the modeling uncertainty. Based on engineering judgment and previous work, the uncertainty from model simplifications and omissions are estimated to be uniformly distributed between −20 % to + 30 %.

Regarding measurement uncertainties, recorded vibrations are subject to the inherent variability in walking gaits of the same person and between individuals [55]. Based on prior measurements, variability in walking gaits has been evaluated for an occupant walking on the same footstep impact locations multiple times and wearing hard- and soft-soled shoes (not included in the evaluation of occupant-tracking strategy). This evaluation results in a measurement uncertainty bounded by the interval of ±45 %, which defines the 99th percentile of the resulting distribution. Based on combined uncertainties, using Monte-Carlo sampling, and target reliability of localization of 95 %, thresholds for each captured footstep event are estimated for occupant localization.

7.2.3. Single-occupant trajectories

Tracking of multiple occupants, walking individually on a full-scale floor slab, is achieved using a model-based approach, as explained in **Section 5**. Examples of CLSs of a few events (intermediate and last events) resulting from occupants walking along trajectory configurations #1 to #6 (see Table 11) are shown in Fig. 13.

In Fig. 13, CLSs are obtained using EDMF and processed using sequential analysis and trajectory-determination operation. CLSs associated with trajectories #2, #3 #5 and #6 in Fig. 13, are from Occupant O1 walking at a moderate speed level (~1.6 Hz) and wears hard-soled shoes during measurement. In Fig. 13, CLSs associated with Trajectory #1 are from Occupant O3 walking with a fast speed level (~1.8 Hz) and wears intermediate-soled shoes while CLSs associated with Trajectory #4 are from Occupant O3 walking with a moderate speed level.

In Fig. 13, small green squares represent the CLSs, and dots represent the falsified location sets. Dashed lines represent the separation walls, and large black squares represent the concrete columns. Diamonds represent sensor locations, and real footstep locations of occupants are represented with crosses.

In Fig. 13, CLSs provided by EDMF always contain the true impact locations, as represented by crosses. Incorporating physics-based models and systematic errors in the interpretation of event signals from single occupants using EDMF, thus provides accurate localization results (accuracy of 100 %) for all footstep events associated with all trajectories.

Moreover, a sequential analysis that accounts for information from previous events, and trajectory determination that is constrained by the condition that occupants walk without backtracking resulted in precise localization. Falsified locations associated with all footstep events, in Fig. 13 present more than 80 % of the initial location set (796 possible locations). Also, in Fig. 13 model-based occupant tracking has led to accurate and precise determination of correct arrival points for all tests once all detected footstep events are investigated.

Although model-based occupant tracking results in precise localization, ambiguous CLs associated with several detected events (for instance, Footstep #27 for Trajectory #1 and Footstep #23 for Trajectory #4 in Fig. 13) remain unfalsified. This is due to the implementation of EDMF to incorporate systematic errors for accurate identification while sacrificing precision. A more comprehensive finite-element model that accounts for the separation walls may improve precision of localization results.

The average accuracy and precision of localization results that are obtained from the model-based occupant tracking are presented in Table 12. For each trajectory (trajectories #1 to #6 in Table 11), the average accuracy and precision are determined based on CLSs resulting from all captured footstep events (see Table 12) repeated multiple times by the three occupants O1, O2 and O3 (see the number of tests in Table 12).

In Table 12, the localization accuracy is determined based on comparing the true footstep locations with the resulting CLSs with a tolerance of plus or minus two footstep locations. This tolerance is taken since the impact locations do not necessarily coincide with the initial location set (796 possible locations) and occupants walk with a self-selected step-length during measurements. Thus, for each footstep event, localization is accurate when at least one candidate location is within 1.5 m radius from the correct location. The radius is equal to twice the distance between two footsteps. Localization precision refers to the percentage of falsified locations from all possible locations (796 possible locations).

Regarding all measurement tests following trajectories #1 to #6, the average localization accuracy using the occupant-tracking strategy varies approximately between 91 % and 97 % as shown in Table 12. Moreover, the average accuracy for all walking tests is approximately 95 %. Thus, the model-based occupant tracking strategy can be applied to multiple occupants walking individually while wearing various types of shoes at self-selected step-length and walking speed to identify accurately their locations within a large full-scale floor slab.

In addition, the average localization precision, regarding all measurement tests following trajectories #1 to #6, varies approximately between 89 % and 93 %, as shown in Table 12. The average precision for all walking tests is approximately 91 %. Thus, incorporating physics-based models in the interpretation of measured footstep events using EDMF and accounting for information from previously detected events using sequential analysis and trajectory determination operations has the potential to provide precise localization results.

Assuming that all departure/arrival points (see X1 to X12 in Fig. 12) are predefined, the model-based occupant-tracking operation provides accurate and precise candidate trajectories for all tested measurements, as presented in Table 13. A summary of the range of the number of candidate trajectories at the last detected footstep event for each tested trajectory (see trajectory configurations #1 to #6 in Table 11) is presented in Table 13. In Table 13, each trajectory is repeated multiple times, back and forth by occupants O1, O2 and O3.

For all tests along each trajectory, the average tracking accuracy varies between 88 % and 100 % in which the correct paths among 132 possible trajectories are determined once all detected events are investigated as shown in Table 13. For all tested trajectories (80 walking tests of single occupants), the model-based tracking strategy has revealed an average accuracy of approximately 93 %.

From Table 13, tracking precision varies between 99 % and 100 % for all walking tests along each trajectory. Tracking precision refers to the percentage of falsified trajectories from all possible paths (132 possible trajectories). when only one possible trajectory remains unfalsified, then occupant tracking is taken to be 100 % precise. No more than 4 candidate trajectories are determined from all tests. The average tracking precision is approximately 99 %.

Therefore, combining the knowledge of structural behavior with measurements and taking into account various sources of uncertainties, model-based occupant tracking provides accurate and precise candidate trajectories of multiple occupants having a range of weights and shoe types walking individually with a self-selected step-length and speed level on a large full-scale floor slab (see Fig. 12).

7.2.4. Two-occupant trajectories

Examples of CLSs of events (intermediate and last events) resulting from two occupants walking along with three trajectory configurations (#9, #10 and #11 in Table 11) are shown in Fig. 14. As explained in Table 11, following trajectory configuration #9, both occupants walk along with Trajectory #3 (see Fig. 12) in opposite directions. For the trajectory configuration #10, both occupants walk along with the same trajectory (see Trajectory #5 in Fig. 12) with an average spacing of four-footsteps. For the trajectory configuration #11, one occupant walks along with Trajectory #1 starting from departure point X3 (see Fig. 12), while the other occupant walks along with Trajectory #3 starting from departure point X1 (see Fig. 12).

For the footstep locations in Fig. 14, occupant localization using EDMF has led to accurate localization results for all detected events from two occupants. Moreover, a sequential analysis that accounts for information from previous events, and trajectory determination have proved to increase localization precision. It can be observed in Fig. 14 that the falsified location sets of the last few events are significant compared with the first few events and this is achieved without compromising accuracy.

In Fig. 14, model-based occupant tracking is shown to lead to accurate determination of correct arrival points for all tested trajectories. During tests, occupants have remained immobile upon reaching their destinations. For the trajectory configuration #11 (see Table 11), the first occupant requires 33 footstep events to attain the arrival point X4 from X1 (see Fig. 12), while the second occupant requires 28 footstep events to attain the arrival point X1 from X3 (see Fig. 12). Thus, starting from the 29th event, model instances related to the arrival point X1 are entirely falsified, since contributions to floor responses are only from the first occupant, as illustrated in Fig. 14. For trajectory configuration #11, the walking path, X1 to X3 is achieved with an average precision of 83 % (25 out of 132 possible trajectories), while the walking path X4 to X1 is determined precisely after exploring all detected events.

EDMF, by including structural information and taking into account systematic errors and model bias, accurately localizes two-occupant locations in a full-scale structure as shown in Fig. 14. However, several ambiguities in the interpretation of measured floor response remain in the resulting CLSs, as shown in Fig. 14. This may be due to the omission of the separation walls in the finite element model and due to the employment of an idealized footstep-impact load function. A sensitivity analysis to evaluate the influence of the separation walls on the simulated responses, is required.

The model-based occupant-tracking operation provides precise and moderately accurate candidate trajectories for two occupants walking simultaneously, as presented in Table 14. Trajectory determination is taken to be accurate when the correct paths among 132 possible trajectories are determined once all detected events are investigated. Tracking precision refers to the percentage of rejected trajectories from all possible paths (132 possible trajectories). If only one possible trajectory per occupant remains unfalsified after investigating the required average footstep events to reach the destination, then occupant tracking is taken to be 100 % precise. For

Table 12
Accuracy and precision of occupant localization for all walking tests following trajectories #1 to #6 (see Fig. 12 and Table 11).

Trajectory	Number of tests	Average accuracy (%)	Average precision (%)
1	20	91.3	91.3
2	10	97.0	91.3
3	12	99.0	91.1
4	16	91.1	93.1
5	10	94.1	89.2
6	12	95.3	91.2
	Average (%)	**94.6**	**91.2**

Table 13
Single-occupant tracking results. Tracking accuracy and precision are for all walking tests following trajectories #1 to #6 (see Fig. 12 and Table 11).

Trajectory configuration	Number of tests	Number of candidate trajectories at the last event	Average accuracy (%)	Average precision (%)
1	20	1–3	90	99.3
2	10	1–2	90	99.9
3	12	1–3	100	99.7
4	16	1–4	87.5	98.9
5	10	1–2	100	99.8
6	12	1–3	91.7	99.0
		Average (%)	**93.2**	**99.4**

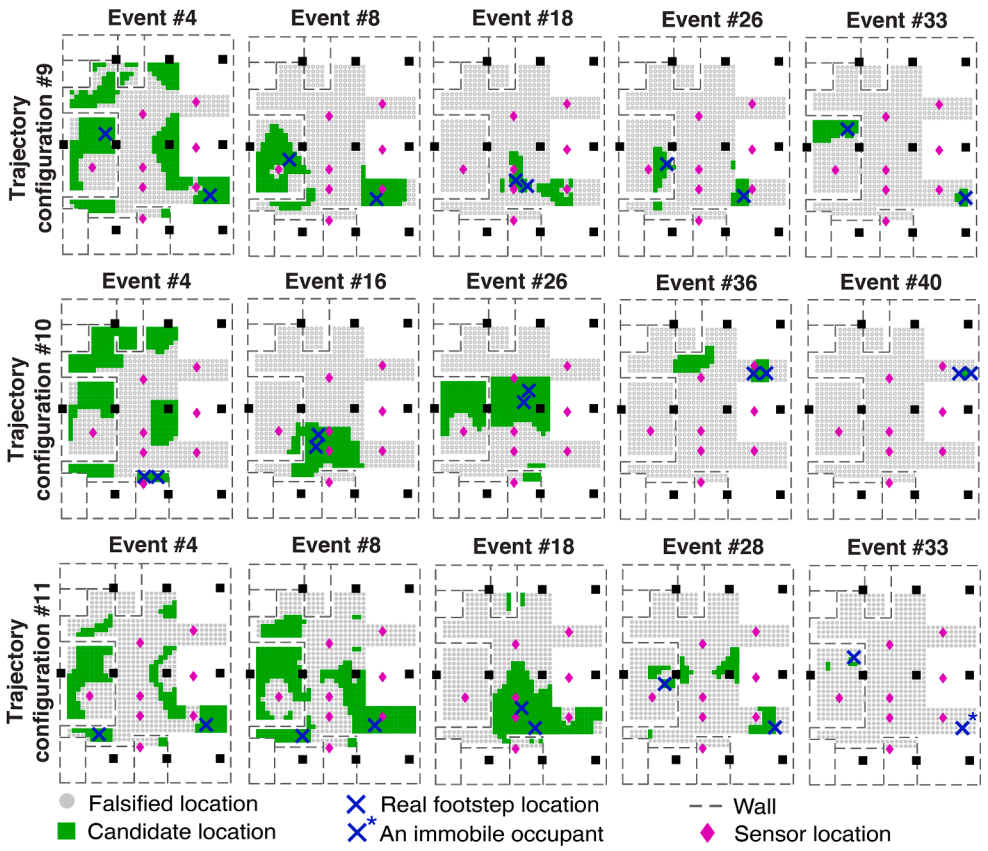

Fig. 14. CLSs that result from two occupants walking along with trajectory configurations #9, #10 and #11 (see Fig. 12 and Table 11) are obtained using EDMF, sequential analysis and trajectory determination.

each trajectory configuration, the average precision is evaluated only for the tests that provide accurate candidate trajectories.

An average tracking precision of 99.5 % is observed for all measurements that reveal accurate tracking. For trajectory configurations #9, 10# and #12, both occupants have carried out approximately the same number of footstep events to reach their destinations (33, 32 and 40 detected events) while for the remaining trajectory configurations #7, #8 and #11 results in a different number of footstep events carried out from each occupant.

In addition, model-based occupant tracking results in accurate trajectory determination for occupants following trajectory configurations #7, #9 and #12 for which the average accuracy from all measurements varies between 83.3 % and 100 % (see Table 14). However, for the remaining trajectory configurations #8, #10 and #11, occupant tracking results in low accuracy rates that vary between 33.3 % and 66.6 % (see Table 14). The average tracking accuracy for all trajectory configurations is found to be 69.4 %.

Occupant tracking using sequential analysis and trajectory determination provided precise tracking of two occupants, as shown in Table 14. In addition, model-based occupant tracking has the potential to accurately determine occupant trajectories. As shown in Table 14, low tracking accuracy is observed for several walking tests. This is due to the recorded footstep signals from two occupants are found to be either off-synchronized or staggered (see **Section 5.1**), leading to capture of more event signals for a trajectory compared with those observed from a single occupant. This induces ambiguities in the resulting CLSs from the localization process that make the continuity of a path from a departure point hard to track. Also, trajectory determination may result in incorrect tracking

Table 14
Two-occupant tracking results. Tracking accuracy and precision are for all walking tests following trajectory configurations #7 to #12 (see Fig. 12 and Table 11).

Trajectory configuration	Number of tests	Number of candidate trajectories at the last event	Average accuracy (%)	Average precision (%)
7	6	2 – 5	83.3	98.5
8	6	1 – 2	33.3	100
9	6	2 – 3	83.3	99.8
10	9	1 – 2	66.6	99.7
11	6	2 – 3	50	99.2
12	6	1	100	100
		Average (%)	**69.4**	**99.5**

when two occupants starting from distant departures are in close proximity. This may generate confusion in the continuity of each trajectory.

Moreover, the number of model instances increases exponentially when detecting multiple occupants walking simultaneously, which may limit the applicability of model-based approaches for crowded spaces. Signal processing techniques to separate emission sources may be useful to localize more than two occupants walking together. Such work is further discussed in the next section.

8. Discussion and limitations

A comprehensive framework for automatic detection and tracking of building occupants based on recorded footstep-induced floor vibrations is successfully evaluated on full-scale varying-rigidity floors (see Case Study 1 in **Section 6**) that are instrumented with sparse sensor configurations (one sensor per ~ 12 to 75 m^2). The framework combines model-free approaches for occupancy detection and counting, with structural-behavior model falsification for tracking.

Accurate event detection is achieved through assessment of the maximum standard deviation from all sensors ($STD_{max,f}$) over segmented and decomposed vibration measurements at multiple frequency ranges that cover the fundamental vertical modes of the structure. An event is detected when a local maximum resulting from $STD_{max,f}$ values exceeds a fixed detection threshold (DT_f) over at least one decomposed signal indicate the occurrence times of a possible event. The number of frequency ranges that cover the fundamental vertical modes of the structure is fixed based on prior measurement observations and engineering judgement. A more comprehensive modal analysis would better define the informative frequency ranges for event detection.

In addition, the level of ambient floor vibrations in terms of magnitudes may vary due to external factors, such as outside traffic that may generate stationary vibrations. Also, internal factors such as activated devices that may operate at low frequency ranges can affect the level of the ambient vibrations. The proposed strategy involves the assumption of a constant level of ambient vibrations, which may not reflect reality. Not accounting for such variations in the level of ambient vibrations might lead to inaccurate detection of events. Thus, automatic quantification of ambient vibrations to update the detection thresholds in real time would increase the reliability of event-detection.

Classification of events into footstep and non-footstep events is successfully carried out using a binary-SVM. Time-and-frequency domain features are used for training. Time-domains features are assessed for event signals decomposed and reconstructed at low-and-high frequency ranges. These ranges cover the frequency band that sensors provide. Null-hypothesis tests are used to select the most informative frequency ranges of time-domain features in order to improve the efficiency of the event classification methodology.

However, the event-classification strategy has been tested only on a single case study using limited sources of non-footstep events from (chair-dragging, opening/closing-door, book-dropping, and hand and mug impacts on a table). In real applications, several non-footstep events from other sources such as electrical devices, falling objects and opening/closing-windows and drawers can occur. Non-included spurious events within the training set and which may have similar signatures to footstep events may decrease the classification performance. More non-footstep-event signals from other sources such as falling objects and opening/closing-windows and drawers could be considered in the training set. In addition, a performance-based analysis of the classifier regarding the most appropriate ratio of footstep events to spurious events is needed.

Footstep-induced floor vibrations are affected by the structural behavior of the floor slab. Change in floor characteristics leads to a change in magnitudes of the footstep-induced vibrations. Therefore, application of the proposed classification strategy to another floor slab requires a commissioning phase to study the structural characteristics to determine the frequency ranges that define the time-domain metrics. Also, re-training the learning algorithm with appropriate vibration measurements is required. Thus, further evaluation of event classification using additional cases of full-scale floor slabs would improve performance.

Event classification could be enhanced through the use of the selected frequency ranges of time-domain features, as explained in **Section 4.2**, to train a one-class SVM. The training data for the one-class SVM classifier is limited to footstep events only (no spurious events). This could improve the event classification strategies proposed by Lam et al. [22] and Pan et al. [67].

Counting the number of occupants has been achieved using an SVM classifier. However, the counting classifier results in low-performance score for two, three, four and five occupants. The determination of two occupants walking together on the floor may be confused with one occupant. Three occupants may be confused with one and two occupants, and finally the presence of four occupants is confused with five occupants. These low performance scores may by due to the low number of events induced by two, three, four and five occupants walking together within the training set compared with those induced by single occupants (7 times more).

Accurate occupant localization is carried out using EDMF on full-scale floor slabs of Case Studies 1 and 2 (see **Sections 6 and 7**).

Incorporating the knowledge of structural behavior through physics-based models in the interpretation of vibration measurements to identify occupant locations allows the use of significantly more sparse sensor configurations on varying-rigidity floor slabs when compared with data-driven techniques [40,44].

Major challenges in data interpretation for occupant localization are related to the generation of reliable footstep-impact simulations as well as the estimation of systematic errors from multiple sources. Model predictions are carried out by dividing the finite element model of the floor slab into a grid of possible locations. Then, the footstep-impact load model is applied separately at each possible location. This excludes the contribution of the other foot during the pre-swing phase of the gait pattern during which the two feet are in contact with the ground. This further increases the uncertainty associated with the simulation model. Thus, accounting for the spacing between succeeding footsteps when walking could improve footstep-impact simulations. Also, improving simulation models, the footstep load model, and the estimation of uncertainties may increase performance in situations where populations of locations are large.

The number of model instances increases exponentially with the number of occupants walking simultaneously, which may limit the applicability of model-based approaches. Signal-processing techniques to separate emission sources may be useful to localize more than two occupants walking together. Several methods including blind source separation (BSS) [93] and equivariant adaptive separation (EAS) [94] methods have the potential to separate overlapping signals. Studying the value of information on the number of occupants to be localized while walking simultaneously is also needed in a range of practical contexts.

Low tracking accuracy is observed for several walking tests (see **Section 7.2.4**). This is due to the recorded footstep signals from two occupants that are found to be either off-synchronized or staggered (see **Section 2**). it is observed that off-synchronized and staggered footstep event may lead to generate higher number of events within the recorded vibrations for a tracking compared with those observed from a single occupant. This induces ambiguities in the resulting CLSs from the localization process that make the continuity of a path from a departure point hard to track. Also, trajectory determination may result in incorrect tracking when two occupants starting from distant departures are in close proximity. This may generate ambiguities in the continuity of each trajectory.

Careful placement of sensors over the floor slab is necessary to guarantee good performance of classifiers and algorithms for occupant detection and tracking. Determining the most informative sensor configuration could be carried out using the joint entropy of simulated footstep impacts. Based on multiple sensor locations the joint entropy assesses the information gain of a set of sensor locations while taking into account the mutual information between their locations [95]. Thus, joint entropy can be used as a metric to design efficient measurement systems.

9. Conclusions

The results of Case Study 1 validate strategies for detection as well as tracking of single occupants. For Case Study 2, the model-based tracking strategy is again validated for single occupants and further validated for two occupants walking together. The following conclusions are drawn:

- A comprehensive framework for detection and tracking of building occupants has been successfully validated using full-scale case studies. The framework involves model-free approaches for occupant detection and a model-based approach for occupant tracking.
- Combining information from multiple frequency components of measurements improves the accuracy of event detection.
- Selection of appropriate frequency components for training enhances the performance of classifiers that distinguish between footstep events (single and multiple occupants walking simultaneously) and non-footstep events.
- Using cross-correlations between measurements at several sensor locations improves the performance of the classifier to distinguish between the presence of either one, two, three, four or five occupants.
- Model-based identification (using EDMF), that includes structural information and takes into account systematic errors and model bias, is able to accurately localize single and two occupants walking simultaneously in full-scale structures. Occupants walk with self-selected step lengths, speed levels and shoe types.
- Occupant tracking using sequential analysis and trajectory determination provides accurate and precise trajectories for up to two occupants walking simultaneously.

Funding

This work was funded by the applied computing and mechanics laboratory EPFL and the Singapore-ETH center (SEC) under contract no. FI 370074011–370074016.

Author contributions

Vibration measurements induced by ambient noise, footstep (from multiple people), and non-footstep impacts were carried out by Slah DRIRA (SD). Analysis of floor vibrations was performed by SD. SD developed occupant detection (event classification and occupant counting) using supervised learning algorithms. SD also developed a strategy for occupant localization and tracking using a model-based data-interpretation approach and created the footstep-impact load model for simulations. Ian SMITH (IS) assisted throughout the elaboration of the methodology and measurement interpretation. IS was actively involved in developing and adapting the data-interpretation methodology using error-domain model falsification (EDMF). IS was involved in interpreting occupant detection, localization, and tracking results. All authors reviewed and accepted the final version of the manuscript.

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.

Acknowledgment

The authors acknowledge the participants that have been involved in vibration measurements: Amal Trabelsi, Marco Proverbio, Gennaro Senatore, Kazuki Hayashi, Arka Reksowardojo, Alireza Khodaverdian, Yafeng Wang and Yves Reuland. The authors also acknowledge Tectus Dreamlab Pte Ltd and BBR Holdings Singapore for access and resources provided for full-scale evaluations of strategies that are described in this paper.

References

[1] B. Song, H. Choi, H.S. Lee, Surveillance tracking system using passive infrared motion sensors in wireless sensor network, in: 2008 Int. Conf. Inf. Netw., 2008: pp. 1–5.
[2] W.P.L. Cully, S.L. Cotton, W.G. Scanlon, J.B. McQuiston, Localization algorithm performance in ultra low power active RFID based patient tracking, in: IEEE 22nd Int Symp. Pers. Indoor Mob. Radio Commun., 2011, pp. 2158–2162.
[3] W.P.L. Cully, S.L. Cotton, W.G. Scanlon, Empirical performance of RSSI-based Monte Carlo localisation for active RFID patient tracking systems, Int. J. Wirel. Inf. Networks. 19 (3) (2012) 173–184.
[4] G. Diraco, A. Leone, P. Siciliano, People occupancy detection and profiling with 3D depth sensors for building energy management, Energy Build. 92 (2015) 246–266.
[5] C.M. Stoppel, F. Leite, Integrating probabilistic methods for describing occupant presence with building energy simulation models, Energy Build. 68 (2014) 99–107.
[6] A.J. Newman, D.K.C. Yu, D.P. Oulton, New insights into retail space and format planning from customer-tracking data, J. Retail. Consum. Serv. 9 (5) (2002) 253–258.
[7] V. Uotila, P. Skogster, Space management in a DIY store analysing consumer shopping paths with data-tracking devices, Facilities. 25 (9/10) (2007) 363–374.
[8] V.L. Erickson, S. Achleitner, A.E. Cerpa, POEM: Power-efficient occupancy-based energy management system, in: Proc. 12th Int. Conf. Inf. Process. Sens. Networks, Philadelphia, Pennsylvania, USA, 2013: pp. 203–216.
[9] P. Henry, M. Krainin, E. Herbst, X. Ren, D. Fox, RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments, Int. J. Rob. Res. 31 (2012) 647–663.
[10] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, K. Whitehouse, The smart thermostat: using occupancy sensors to save energy in homes, in: Proc. 8th ACM Conf. Embed. Networked Sens. Syst., Zürich, Switzerland, 2010: pp. 211–224.
[11] W. Wang, J. Chen, T. Hong, Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings, Autom. Constr. 94 (2018) 233–243.
[12] J. Yu, P. Wang, T. Koike-Akino, Y. Wang, P. V Orlik, H. Sun, Human Pose and Seat Occupancy Classification with Commercial MMWave WiFi, in: 2020 IEEE Globecom Work. (GC Wkshps, 2020: pp. 1–6.
[13] X. Huang, H. Cheena, A. Thomas, J.K.P. Tsoi, B. Gao, Indoor Detection and Tracking of People Using mmWave Sensor, J. Sensors. 2021 (2021) 1–14.
[14] Y. Zeng, P.H. Pathak, P. Mohapatra, WiWho: wifi-based person identification in smart spaces, in: Proc. 15th Int. Conf. Inf. Process. Sens. Networks, 2016: p. 4.
[15] C. Xu, B. Firner, R.S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, N. An, SCPL: indoor device-free multi-subject counting and localization using radio signal strength, in: Proc. 12th Int. Conf. Inf. Process. Sens. Networks, Philadelphia, PA, USA, 2013: pp. 79–90.
[16] N. Li, B. Becerik-Gerber, Performance-based evaluation of RFID-based indoor location sensing solutions for the built environment, Adv. Eng. Informatics. 25 (3) (2011) 535–546.
[17] M. Han, J. Zhao, X. Zhang, J. Shen, Y. Li, The reinforcement learning method for occupant behavior in building control: a review, Energy, Built Environ. 2 (2) (2021) 137–148.
[18] C.X. Lu, S. Rosa, P. Zhao, B. Wang, C. Chen, J.A. Stankovic, N. Trigoni, A. Markham, See through smoke: robust indoor mapping with low-cost mmwave radar, in, Proc. 18th Int. Conf. Mob. Syst. Appl. Serv. (2020) 14–27.
[19] P. Lazik, N. Rajagopal, O. Shih, B. Sinopoli, A. Rowe, ALPS: A bluetooth and ultrasound platform for mapping and localization, in: Proc. 13th ACM Conf. Embed. Networked Sens. Syst., Seoul, South Korea, 2015: pp. 73–84.
[20] J.T. Biehl, M. Cooper, G. Filby, S. Kratz, Loco: a ready-to-deploy framework for efficient room localization using wi-fi, in: Proc. 2014 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput., 2014: pp. 183–187.
[21] S. Pan, A. Bonde, J. Jing, L. Zhang, P. Zhang, H.Y. Noh, Boes: building occupancy estimation system using sparse ambient vibration monitoring, in: Sensors Smart Struct. Technol. Civil, Mech. Aerosp. Syst. 2014, San Diego, California, United States, 2014: p. 90611O.
[22] M. Lam, M. Mirshekari, S. Pan, P. Zhang, H.Y. Noh, Robust occupant detection through step-induced floor vibration by incorporating structural characteristics, in: Dyn. Coupled Struct. Vol. 4, Springer, 2016: pp. 357–367.
[23] M. Mirshekari, S. Pan, P. Zhang, H.Y. Noh, Characterizing wave propagation to improve indoor step-level person localization using floor vibration, in: Sensors Smart Struct. Technol. Civil, Mech. Aerosp. Syst. 2016, Las Vegas, Nevada, USA, 2016: p. 980305.
[24] S. Pan, T. Yu, M. Mirshekari, J. Fagert, A. Bonde, O.J. Mengshoel, H.Y. Noh, P. Zhang, https://doi.org/10.1145/3130954, in: Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol., ACM, 2017: pp. 1–31. https://doi.org/10.1145/3130954.
[25] S. Narayana, R.V. Prasad, V.S. Rao, T. V Prabhakar, S.S. Kowshik, M.S. Iyer, PIR sensors: Characterization and novel localization technique, in: Proc. 14th Int. Conf. Inf. Process. Sens. Networks, Seattle, Washington, 2015: pp. 142–153.
[26] W.-S. Zheng, S. Gong, T. Xiang, Person re-identification by probabilistic relative distance comparison, in, CVPR 2011 (2011) 649–656.
[27] S. Budi, K. Hyoungseop, T.J. Kooi, I. Seiji, Real time tracking and identification of moving persons by using a camera in outdoor environment, (2009).
[28] D.T. Alpert, M. Allen, Acoustic gait recognition on a staircase, in: 2010 World Autom. Congr., 2010: pp. 1–6.
[29] J.T. Geiger, M. Kneißl, B.W. Schuller, G. Rigoll, Acoustic gait-based person identification using hidden Markov models, in: Proc. 2014 Work. Mapp. Personal. Trait. Chall. Work., 2014: pp. 25–30.
[30] L.M. Candanedo, Véronique Feldheim, Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models, Energy Build. 112 (2016) 28–39.
[31] C. Jiang, M.K. Masood, Y.C. Soh, H. Li, Indoor occupancy estimation from carbon dioxide concentration, Energy Build. 131 (2016) 132–141.
[32] R. Serra, P. Di Croce, R. Peres, D. Knittel, Human step detection from a piezoelectric polymer floor sensor using normalization algorithms, in: SENSORS, 2014 IEEE, 2014: pp. 1169–1172.
[33] R. Serra, D. Knittel, P. Di Croce, R. Peres, Activity recognition with smart polymer floor sensor: Application to human footstep recognition, IEEE Sens. J. 16 (2016) 5757–5775.
[34] Z.D. Tekler, R. Low, B. Gunay, R.K. Andersen, L. Blessing, A scalable Bluetooth Low Energy approach to identify occupancy patterns and profiles in office spaces, Build. Environ. 171 (2020), 106681.

[35] K. Weekly, N. Bekiaris-Liberis, M. Jin, A.M. Bayen, Modeling and estimation of the humans' effect on the CO 2 dynamics inside a conference room, IEEE Trans. Control Syst. Technol. 23 (2015) 1770–1781.
[36] S. Drira, Y. Reuland, S.G.S. Pai, H.Y. Noh, I.F.C. Smith, Model-Based Occupant Tracking Using Slab-Vibration Measurements, Front. Built Environ. 5 (2019) 63, https://doi.org/10.3389/fbuil.2019.00063.
[37] S. Drira, Y. Reuland, N.F.H. Olsen, S.G.S. Pai, I.F.C. Smith, Occupant-detection strategy using footstep-induced floor vibrations, in: Proc. 1st ACM Int. Work. Device-Free Hum. Sens., ACM, New York, NY, USA, 2019: pp. 31–34. https://doi.org/10.1145/3360773.3360881.
[38] S.G.S. Pai, Y. Reuland, S. Drira, I.F.C. Smith, Is there a relationship between footstep-impact locations and measured signal characteristics?, in: 1st ACM Int. Work. Device-Free Hum. Sens., New York, USA, 2019.
[39] S. Drira, Y. Reuland, I.F.C. Smith, Model-based interpretation of floor vibrations for indoor occupant tracking, in: 26th Int. Work. Intell. Comput. Eng., Leuven Belgium, 2019.
[40] M. Mirshekari, S. Pan, J. Fagert, E.M. Schooler, P. Zhang, H.Y. Noh, Occupant localization using footstep-induced structural vibration, Mech. Syst. Signal Process. 112 (2018) 77–97.
[41] S.A. Alyamkin, S.I. Eremenko, Pedestrian detection algorithms based on an analysis of the autocorrelation function of a seismic signal, Optoelectron. Instrum. Data Process. 47 (2) (2011) 124–129.
[42] A. Subramanian, K.G. Mehrotra, C.K. Mohan, P.K. Varshney, T. Damarla, Feature selection and occupancy classification using seismic sensors, in: Int. Conf. Ind. Eng. Other Appl. Appl. Intell. Syst., 2010: pp. 605–614.
[43] R.K. Begg, M. Palaniswami, B. Owen, Support vector machines for automated gait classification, IEEE Trans. Biomed. Eng. 52 (5) (2005) 828–838.
[44] J. Clemente, F. Li, M. Valero, W. Song, Smart seismic sensing for indoor fall detection, location and notification, IEEE J. Biomed. Heal. Informatics. (2019).
[45] J. Clemente, W. Song, M. Valero, F. Li, X. Liy, Indoor person identification and fall detection through non-intrusive floor seismic sensing, IEEE Int Conf. Smart Comput. 2019 (2019) 417–424.
[46] Y. Liao, V.R. Vemuri, Use of k-nearest neighbor classifier for intrusion detection, Comput. Secur. 21 (5) (2002) 439–448.
[47] S. Drira, S.G.S. Pai, Y. Reuland, N.F.H. Olsen, I.F.C. Smith, Using footstep-induced vibrations for occupant detection and recognition in buildings, Adv. Eng. Informatics. (2020).
[48] Y. Zhang, S. Pan, J. Fagert, M. Mirshekari, H.Y. Noh, P. Zhang, L. Zhang, Occupant activity level estimation using floor vibration, Proc. 2018 ACM Int. Jt. Conf. 2018 Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput. (2018) 1355–1363.
[49] J.D. Poston, R.M. Buehrer, P.A. Tarazaga, A framework for occupancy tracking in a building via structural dynamics sensing of footstep vibrations, Front. Built Environ. 3 (2017) 65.
[50] S. Pan, M. Mirshekari, P. Zhang, H.Y. Noh, Occupant traffic estimation through structural vibration sensing, in: Sensors Smart Struct. Technol. Civil, Mech. Aerosp. Syst. 2016, Las Vegas, Nevada, USA, 2016: p. 980306.
[51] J.M. Zurada, Introduction to artificial neural systems, West St, Paul, 1992.
[52] J. Schloemann, V.V.N.S. Malladi, A.G. Woolard, J.M. Hamilton, R.M. Buehrer, P.A. Tarazaga, Vibration event localization in an instrumented building, in: Exp. Tech. Rotating Mach. Acoust. Vol. 8, Springer, 2015: pp. 265–271.
[53] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, H.Y. Noh, Obstruction-invariant occupant localization using footstep-induced structural vibrations, Mech. Syst. Signal Process. 153 (2021), 107499.
[54] S.G.S. Pai, I.F.C. Smith, Multi-fidelity modelling for structural identification, in: Towar. a Resilient Built Environ. IABSE Symp. Guimaraes 2019, 2019.
[55] S. Drira, S.G.S. Pai, I.F.C. Smith, Uncertainties in structural behavior for model-based occupant localization using floor vibrations, Front. Built Environ. 7 (2021) 13.
[56] J. Fagert, M. Mirshekari, S. Pan, L. Lowes, M. Iammarino, P. Zhang, H.Y. Noh, Structure-and Sampling-Adaptive Gait Balance Symmetry Estimation Using Footstep-Induced Structural Floor Vibrations, J. Eng. Mech. 147 (2021) 4020151.
[57] Y.H. Lee, T. Oh, The measurement of P-, S-, and R-wave velocities to evaluate the condition of reinforced and prestressed concrete slabs, Adv. Mater. Sci. Eng. 2016 (2016) 1–14.
[58] R. Bahroun, O. Michel, F. Frassati, M. Carmona, J.L. Lacoume, New algorithm for footstep localization using seismic sensors in an indoor environment, J. Sound Vib. 333 (3) (2014) 1046–1066.
[59] W. Chen, M. Guan, L. Wang, R. Ruby, K. Wu, FLoc: Device-free passive indoor localization in complex environments, in,, IEEE Int, Conf. Commun. 2017 (2017) 1–6.
[60] J.D. Poston, J. Schloemann, R.M. Buehrer, V.V.N.S. Malladi, A.G. Woolard, P.A. Tarazaga, Towards indoor localization of pedestrians via smart building vibration sensing, in: 2015 Int. Conf. Locat. GNSS, 2015: pp. 1–6.
[61] S. Pan, K. Lyons, M. Mirshekari, H.Y. Noh, P. Zhang, Multiple pedestrian tracking through ambient structural vibration sensing, in: Proc. 14th ACM Conf. Embed. Netw. Sens. Syst. CD-ROM, 2016: pp. 366–367.
[62] L. Shi, M. Mirshekari, J. Fagert, Y. Chi, H.Y. Noh, P. Zhang, S. Pan, Device-free Multiple People Localization through Floor Vibration, in, Proc. 1st ACM Int. Work. Device-Free Hum. Sens. (2019) 57–61.
[63] M.S. Ford, The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Health Phys. 84 (2003) 667–668.
[64] S. Drira, Y. Reuland, I.F.C. Smith, Occupant tracking using model-based data interpretation of structural vibrations, in: 9th Int. Conf. Struct. Heal. Monit. Intell. Infrastruct., St. Louis, MO, USA, 2019.
[65] J.-A. Goulet, I.F.C. Smith, Structural identification with systematic errors and unknown uncertainty dependencies, Comput. Struct. 128 (2013) 251–258.
[66] R.D. Gregory, I. Gladwell, The reflection of a symmetric Rayleigh-Lamb wave at the fixed or free edge of a plate, J. Elast. 13 (2) (1983) 185–206.
[67] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H.Y. Noh, P. Zhang, Indoor person identification through footstep induced structural vibration, in: Proc. 16th Int. Work. Mob. Comput. Syst. Appl., Santa Fe, New Mexico, USA, 2015: pp. 81–86.
[68] V. Racic, A. Pavic, J.M.W. Brownjohn, Experimental identification and analytical modelling of human walking forces: Literature review, J. Sound Vib. 326 (1-2) (2009) 1–49.
[69] H. Lee, J.W. Park, A.S. Helal, Estimation of indoor physical activity level based on footstep vibration signal measured by MEMS accelerometer in smart home environments, in, Int. Work. Mob. Entity Localization Track. GPS-Less Environ. (2009) 148–162.
[70] S. Drira, Occupancy detection and tracking in buildings using floor-vibration signals, École Polytechnique Fédérale de Lausanne - EPFL, Thesis n° 8289, 2020.
[71] K. Kanazawa, K. Hirata, Parametric estimation of the cross-power spectral density, J. Sound Vib. 282 (1-2) (2005) 1–35.
[72] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press (2016).
[73] G. Zhang, X. Wang, Y.-C. Liang, J. Liu, Fast and robust spectrum sensing via Kolmogorov-Smirnov test, IEEE Trans. Commun. 58 (12) (2010) 3410–3416.
[74] J. Hauke, T. Kossowski, Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data, Quaest. Geogr. 30 (2) (2011) 87–93.
[75] S. Drira, S.G.S. Pai, I.F.C. Smith, Increasing occupant localization precision through identification of footstep-contact dynamics, Adv. Eng. Informatics. (2020).
[76] F.P. Figueiredo, J.G.S. da Silva, L.R.O. de Lima, A Parametric Study of Composite Footbridges with Pedestrian Walking Loads, in: Proc. Tenth Int. Conf. Civil, Struct. Environ. Eng. Comput., 2005: p. 85.
[77] T.-C. Pan, X. You, C.L. Lim, Evaluation of floor vibration in a biotechnology laboratory caused by human walking, J. Perform. Constr. Facil. 22 (3) (2008) 122–130.
[78] C.-L. Liu, C.-H. Lee, P.-M. Lin, A fall detection system using k-nearest neighbor classifier, Expert Syst. Appl. 37 (10) (2010) 7174–7181.
[79] S. TAN, An effective refinement strategy for KNN text classifier, Expert Syst. Appl. 30 (2) (2006) 290–298.
[80] B. Wu, R. Nevatia, Cluster boosted tree classifier for multi-view, multi-pose object detection, IEEE 11th Int Conf. Comput. Vis. 2007 (2007) 1–8.
[81] E.-J. Ong, R. Bowden, A boosted classifier tree for hand shape detection, in: Sixth IEEE Int. Conf. Autom. Face Gesture Recognition, 2004. Proceedings., 2004: pp. 889–894.
[82] A.M. APDL, Mechanical applications Theory reference, 2010.
[83] M. Stein, Large sample properties of simulations using Latin hypercube sampling, Technometrics. 29 (2) (1987) 143–151.

[84] S.G.S. Pai, A. Nussbaumer, I.F.C. Smith, Comparing Structural Identification Methodologies for Fatigue Life Prediction of a Highway Bridge, Front. Built Environ. 3 (2018) 73.
[85] Y. Reuland, P. Lestuzzi, I.F.C. Smith, A model-based data-interpretation framework for post-earthquake building assessment with scarce measurement data, Soil Dyn. Earthq. Eng. 116 (2019) 253–263.
[86] A. Pavic, M.R. Willford, G. Appendix, Vibration serviceability of post-tensioned concrete floors, Post-Tensioned Concr, Floors Des. Handb. (2005) 99–107.
[87] M.R. Willford, P. Young, Mim., CEng, A design guide for footfall induced vibration of structures, Concrete Society for The Concrete Centre (2006).
[88] A.L. Smith, S.J. Hicks, P.J. Devine, Design of floors for vibration: A new approach, Steel Construction Institute Ascot, Berkshire, UK, 2007.
[89] H. Bachmann, A.J. Pretlove, J.H. Rainer, Vibrations induced by people, in: Vib. Probl. Struct., Springer, 1995: pp. 1–28.
[90] J. Blanchard, B.L. Davies, J.W. Smith, Design criteria and analysis for dynamic loading of footbridges, in: Proceeding a Symp. Dyn. Behav. Bridg. Transp. Road Res. Lab., Crowthorne, Berkshire, England, 1977: pp. 90–106.
[91] S. Živanović, A. Pavic, P. Reynolds, Vibration serviceability of footbridges under human-induced excitation: a literature review, J. Sound Vib. 279 (1-2) (2005) 1–74.
[92] P. Young, Improved floor vibration prediction methodologies, in: ARUP Vib. Semin., 2001.
[93] S. Makino, S. Araki, R. Mukai, H. Sawada, Audio source separation based on independent component analysis, in: 2004 IEEE Int. Symp. Circuits Syst. (IEEE Cat. No. 04CH37512), 2004: pp. V–V.
[94] J.-F. Cardoso, B.H. Laheld, Equivariant adaptive source separation, IEEE Trans. Signal Process. 44 (12) (1996) 3017–3030.
[95] N.J. Bertola, A. Costa, I.F.C. Smith, Strategy to validate sensor-placement methodologies in the context of sparse measurement in complex urban systems, IEEE Sens. J. 20 (10) (2020) 5501–5509.