

# Early warning signals of failures in building management systems

José Joaquín Mesa-Jiménez<sup>1,2,\*</sup>, Lee Stokes<sup>3</sup>, QingPing Yang<sup>1</sup>, and Valerie Livina<sup>2</sup>

<sup>1</sup> Brunel University London, Kingston Lane, Uxbridge, UK

<sup>2</sup> National Physical Laboratory, Hampton Road, Teddington, UK

<sup>3</sup> City Science, The Generator, Kings Wharf, The Quay, Exeter, UK

Received: 29 December 2020 / Accepted: 22 April 2021

**Abstract.** In the context of sensor data generated by Building Management Systems (BMS), early warning signals are still an unexplored topic. The early detection of anomalies can help preventing malfunctions of key parts of a heating, cooling and air conditioning (HVAC) system that may lead to a range of BMS problems, from important energy waste to fatal errors in the worst case. We analyse early warning signals in BMS sensor data for early failure detection. In this paper, the studied failure is a malfunction of one specific Air Handling Unit (AHU) control system that causes temperature spikes of up to 30 degrees Celsius due to overreaction of the heating and cooling valves in response to an anomalous temperature change caused by the pre-heat coil in winter period in a specific area of a manufacturing facility. For such purpose, variance, lag-1 autocorrelation function (ACF1), power spectrum (PS) and variational autoencoder (VAE) techniques are applied to both univariate and multivariate scenarios. The univariate scenario considers the application of these techniques to the control variable only (the one that displays the failure), whereas the multivariate analysis considers the variables affecting the control variable for the same purpose. Results show that anomalies can be detected up to 32 hours prior to failure, which gives sufficient time to BMS engineers to prevent a failure and therefore, an proactive approach to BMS failures is adopted instead of a reactive one.

**Keywords:** Early warning signals / dynamical systems / tipping point / predictive maintenance

## 1 Introduction

Digital transformation involves changes of key business operations which affect products and processes [1], and data analytics and machine learning play a key role in these processes [2]. In recent years, the development of connectivity and flows of information between devices and sensors provided abundant data. In order to extract value from this data, Early Warning Signals (EWS) can be used to reduce systems downtime, optimise capacity and reduce operational costs.

Building Management Systems (BMS) present advantages for energy control and comfort policy management such as identifying locations of potential energy waste, decreasing equipment operating cost, providing indoor environmental safety and comfort through HVAC system control, as well as controls of water consumption, elevators, etc. This paper focuses on the early detection of failures in the temperature control of zones, where a policy comfort

has been maintained. This temperature variable often displays a failure in heating or cooling system. An example could be a failure in a certain point of the system (fan coil unit, ventilation, heating, etc.) that causes the temperature of a zone to decrease well below its policy comfort.

### 1.1 Fault detection and diagnosis

As a part of this study, we consider that the EWS methodologies applied could be relevant to diagnosis of system errors. For AHU fault diagnosis some previous work has been reported in the literature. Reference [3] describes the application of Artificial Neural Networks (ANNs) to fault diagnosis in an AHU by using residuals of system variables to quantify the dominant symptoms of fault modes of operation. Following the same approach, reference [4] proposed AHU subsystem level fault detection using a General Regression Neural-Network (GRNN), residual generation and fault detection and diagnosis. A novel feature extraction technique to extract temperature and power associated features from high-dimensional and unstructured terminal unit data is presented in [5], to

\* Corresponding author: [joaquinmesajimenez@gmail.com](mailto:joaquinmesajimenez@gmail.com)

diagnose faulty HVAC in an automatic and remote manner. Reference [6] exercises the use of Air handling unit Performance Assessment Rules (APAR). They use control signals to determine the mode of operation of the AHU. A subset of expert rules which correspond to that mode is then evaluated to determine whether a fault exists. In the review of fault detection and diagnosis methodologies carried by [7], various Fault Detection and Diagnosis (FDD) are described to illustrate the use of evaluation standard parameters for improving the performance of AHUs. This work divides FDDs in three main categories, namely analytical-based methods, knowledge-based methods, and data-driven methods. In [8] an approach for clustering air conditioning zones of influence together, shows how room areas, weather conditions, and air conditioning settings affect the air conditioning power consumption of rooms in real life. In the context of building's energy performance, reference [9] proposes a hybrid and multilevel FDD tool for the identification and prioritization of corrective maintenance actions helping to ensure the energy performance of buildings using dynamic Bayesian networks to monitor energy consumption. Following a similar line of work in the context of building energy performance, reference [10] proposed to evaluate the uncertainty associated to the use of a simplified model for the estimation of the energy consumption of a given building. In a more recent study, reference [11] proposes a method that employs sequential two-state clustering to identify abnormal behaviour of the fan coil unit. Some other recent studies on HVAC systems fault detection and diagnosis can be seen in [5,12,13].

## 1.2 Early warning signals

The main problem in the state-of-the-art BMSs is that they tend to use a reactive approach instead of a proactive one. The system failure is detected once it has happened, thus causing a disruption in services and forcing engineers to temporarily shut down some of the equipment in order to fix the failure. This paper proposes a proactive approach to early detection of BMS failures, aiming at both EWS and forecast of time series temperature sensor data.

Prior publications prove that EWS have a wide range of applications such as climate pattern change analysis [14–17]. Also, a common application for EWS in manufacturing is degradation assessment [18,19], and early detection of failures for critical system components in general as explained in [20]. Another popular area of application is economics, such as the analysis of banking system collapse in [21–23], credit risk diagnosis [24,25] and studies for economic cycles for certain asset prices in [26–28]. Also in biology, with a wide variety of applications within the field such as [29–32].

There are several studies of EWS in dynamical systems such as climatic variables that can be applied to our case study due to its nature as dynamical system. An example is [14], which uses lag-1 autocorrelation (ACF1), the Detrended Fluctuation Analysis (DFA) exponent and Power Spectrum (PS) in tropical cyclon data. For the

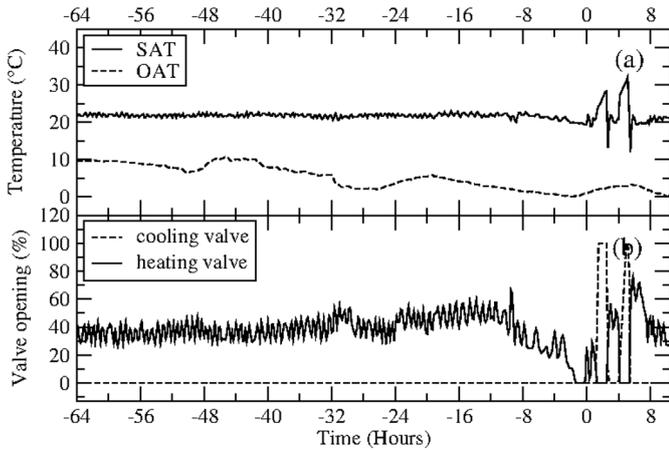
multiivariate tropical cyclone data case, reference [15] uses Empirical Orthogonal Functions (EOF) for dimensionality reduction prior to applying ACF1, DFA and PS. They also study the possibility of using the Jacobian matrix eigenvalues of the system as tipping point indicator. In the same area of application, references [16,17] use ACF and DFA to detect climate tipping points. In [33] a novel statistical method is applied, which is the method of potentials [34], for EWS to analyse the changing number of climate states during the last 60 kyr. The method detects the changes between states by estimating the probability density of the recorded time series. In the predictive maintenance field, reference [35] uses DFA and ACF to detect anomalies in electronic components commonly used in applications of the automotive and aviation industries. A different approach is shown in [36], which uses invariant-based identification for Lithium-ion battery performance degradation. In [37], they present data driven based models for air source heat pumps performance evaluation and anomaly detection based on real data collected over a water heating system.

Variational Autoencoders (VAE) are becoming increasingly popular for failure detection. In [38] they use an autoencoder residual vector error magnitude. This method is tested on several images datasets, concluding that it is a valid methodology for failure detection, as the hidden layer representation is capable of characterising the fundamental attributes of the system within normal conditions, therefore to measure the deviation from “normal functioning”. Following a similar method in a manufacturing field, in [18] uses VAE for degradation assessment of the ball screw. The assessment is done using the Variational Autoencoder Reconstruction Error (VAERE) and it demonstrates the progressive degradation of this component.

The paper is organised as follows. Section 2 describes the data used in this study, both for the univariate and multivariate cases. In Section 3, the methodologies used for the purpose of this paper are explained and summarised. Results are then outlined in Section 4. Finally, Section 5 outlines the conclusions.

## 2 Problem description & data

The goal of this paper is to apply EWS techniques for early failure detection in Building Management Systems (BMS). The building used in this case study is a large manufacturing facility that consists of office spaces and production spaces, separated in different plants. The failure consists of a malfunction of the heating and cooling valve control system that causes unnecessary temperature spikes in the affected zone of the building. When the Outside Air Temperature (OAT) drops below 5 degrees Celsius, the control system activates the frost coil valve, which is designed to protect the HVAC equipment from freezing but also may have an effect in the room temperature (in this case it does). This causes the cooling and heating valve to overreact in order to compensate de influence in temperature and thus, causing the spikes.



**Fig. 1.** Plot of sensor variables affecting room temperature. (a) Supply Air Temperature (SAT) and Outside Air Temperature (OAT) in degrees Celsius. (b) Percentage of opening for each of the heating and cooling valves. Plot window starts at 5:30 AM of the 25th of January 2019, 64 hours prior to the failure.

This malfunction has been observed after a long period of time with no failures.

Our variable of interest is the supply air temperature (SAT) of a particular zone of the building in the manufacturing plant. SAT is controlled by the average room temperature measured by four sensors installed in the zone using the control signal of pre-established fixed setpoint. Average room temperature is then used by the system to regulate the percentage of openness of heating or cooling valves, according to the setpoint. Another factor that influences the temperature of the room is the OAT. These four variables are represented in Figure 1, where the origin, 0, is taken to represent the start of the failure.

Another variable that controls temperature of the room is the frost coil valve (or pre-heat coil), which is worth mentioning as it may have an influence in the room temperature by increasing it, but we are not using it for the purpose of early failure detection.

- Outside Air Temperature controls frost coil valve. Frost coil valve protects the Air Handling Unit (AHU) from very low temperatures (the pre-heat coil protects the parts containing liquid elements, otherwise they would contract due to the low temperatures and damage the AHU as a consequence).
- Supply Air Temperature controls heating and cooling valves. As data is registered in winter, the cooling valve will be closed most of the time, but in periods of anomaly it activates and provides with cooling to the environment.

We first conduct univariate analysis on the SAT by using different indicators explained in Section 3, and then we take into account other variables, reducing their dimensionality and applying the same techniques to compare both approaches and identify which one provides an EWS the earliest.

## 3 Methodology

### 3.1 Early warning signals

There are a number of techniques to detect Early Warning Signals (EWS) as mentioned in Section 1. For the purpose of early failure detection applied to BMS we use four: variance, power spectrum, ACF1 and VAE. These early warning indicators are used with a chosen sliding window applied to the time series preceding the onset of a transition. In this case, the choice of the length of the sliding window is a trade-off between time-resolution (data availability) and the clarity of the change of the signal prior to transition. In order to identify trends before the transition, we use Kendall  $\tau$  correlation coefficient [39] in one of the indicators, PS, as this indicator presents a more chaotic behavior prior to transition. A positive Kendall  $\tau$  coefficient indicates increasing trends in the indicators prior to transition, as applied in [16].

#### 3.1.1 Variance

The first indicator we use is the variance of time series, which is defined by:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}, \quad (1)$$

where  $s^2$  is the sample variance,  $x_i$  is the value of one observation,  $\bar{x}$  is the mean value of all observations and  $n$  is the number of observations. The reason to use variance is to compare a more simple approach with other EWS indicators.

#### 3.1.2 Power spectrum (PS)

Another EWS indicator is the power spectrum scaling exponent  $\beta$ , which is calculated by estimating the slope of the power spectrum  $S(f)$  of the data, plotted on logarithmic axes [40] in short term range. Exponent  $\beta$  can be estimated as:

$$S(f) \sim f^{-\beta}, \quad (2)$$

where the power spectrum is approximated by the periodogram, obtained from the absolute value of the fast Fourier transform. Then we obtain  $\beta$  by measuring the slope inside of the frequency range  $10^{-2} \leq f \leq 10^{-1}$ .

#### 3.1.3 Lag-1 auto-correlation function (ACF1)

Autocorrelation function is also used for the purpose of this paper. This function measures correlation of the time series within itself at different time lags. According to [41], the definition of the lag- $k$  autocorrelation function is:

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

**Algorithm 1: Principal Component Analysis (D,  $\alpha$ ), reproduced from [45]:**


---

```

1  $\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$  ; // Compute mean
2  $\mathbf{Z} = \mathbf{D} - \mathbf{1} \cdot \boldsymbol{\mu}^T$  ; // Center data
3  $\boldsymbol{\Sigma} = \frac{1}{n} (\mathbf{Z}^T \mathbf{Z})$  ; // Compute covariance matrix
4  $(\lambda_1, \lambda_2, \dots, \lambda_d) = \text{eigenvalues}(\boldsymbol{\Sigma})$  ; // Compute eigenvalues
5  $\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_d) = \text{eigenvectors}(\boldsymbol{\Sigma})$  ; // Compute eigenvectors
6  $f(r) = \frac{\sum_{i=1}^r \lambda_i}{\sum_{i=1}^d \lambda_i}, \forall r = 1, 2, \dots, d$  ; // Fraction of total variance
7 Choose smallest  $r$  so that  $f(r) \geq \alpha$  ; // Choose dimensionality capturing  $\alpha$  fraction of
   total variance
8  $\mathbf{U}_r = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_r)$  ; // Reduced basis
9  $\mathbf{A} = \{\mathbf{a}_i | \mathbf{a}_i = \mathbf{U}_r^T \mathbf{x}_i, i = 1, \dots, n\}$  ; // Reduced dimensionality data

```

---

We use lag-1 ( $k=1$ ) autocorrelation function (ACF1), the same as in previous studies. It is important to mention the influence of the right choice of  $k$ . If  $k$  is too small, the indicator would respond very quickly to changes in the time series. If  $k$  is too large, the indicator would barely change prior to a significant change in the series.

### 3.1.4 Variational auto-encoder (VAE)

This technique is a novel methodology for EWS. According to [42], the VAE is a network which attempts to represent the input with a PDF instead of several hidden nodes. This is illustrated in [43], with a training set  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_A]^T$  being  $\mathbf{x}$  vectors from time  $t=1$  to time  $t=A$ .

VAE uses a neural network for the probabilistic encoder  $q_\theta(\mathbf{z}|\mathbf{x})$  to approximate the posterior of the generative model  $p_\theta(\mathbf{x}, \mathbf{z})$ . Let the prior over the latent variables be the centered isotropic multivariate Gaussian  $p_\theta(\mathbf{z}) = \mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I})$ . The parameters of the distribution are computed with a fully-connected neural network with a single hidden layer that attempts to reduce the reconstruction error. According to [43], the posterior is approximated with a multivariate Gaussian with a diagonal covariance structure:

$$\log q_\theta(\mathbf{z}|\mathbf{x}^{(i)}) = \log \mathcal{N}(\mathbf{z}; \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{2(i)} \mathbf{I}) \quad (4)$$

here, the mean and standard deviation,  $\boldsymbol{\mu}^{(i)}$  and  $\boldsymbol{\sigma}^{(i)}$  respectively, are the outputs of the encoding half of the network. Concerning the decoder half of the network, samples are computed from the posterior  $\mathbf{z}^{(i,l)} \sim q_\theta(\mathbf{z}|\mathbf{x}^{(i)})$  using  $\mathbf{z}^{(i,l)} = g_\theta(\mathbf{x}^{(i)}, \boldsymbol{\varepsilon}^{(l)}) = \boldsymbol{\mu}^{(i)} + \boldsymbol{\sigma}^{(i)} \odot \boldsymbol{\varepsilon}^{(l)}$  where  $\boldsymbol{\varepsilon}^{(l)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ , with  $\odot$  being an element-wise multiplication. As both prior ( $p_\theta(\mathbf{z})$ ) and posterior ( $q_\theta(\mathbf{z}|\mathbf{x})$ ) are assumed to be Gaussian, the resulting estimator for the model and datapoint  $\mathbf{x}^{(i)}$  is, according to [43]:

$$\begin{aligned} \mathcal{L}(\theta, \phi; \mathbf{x}^{(i)}) &\simeq \frac{1}{2} \sum_{j=1}^J (1 + \log((\sigma_j^{(i)})^2) - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2) \\ &+ \frac{1}{L} \sum_{l=1}^L \log p_\theta(\mathbf{x}^{(i)} | \mathbf{z}^{(i,l)}), \\ \text{where } \mathbf{z}^{(i,l)} &= \boldsymbol{\mu}^{(i)} + \boldsymbol{\sigma}^{(i)} \odot \boldsymbol{\varepsilon}^{(l)} \text{ and } \boldsymbol{\varepsilon}^{(l)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \end{aligned} \quad (5)$$

We first assume a normal functioning state in the system and training the model parameters ( $\boldsymbol{\mu}^{(i)}$  and  $\boldsymbol{\sigma}^{(i)}$ ) with such data. As EWS indicator, we track the reconstruction error, which we define with the root mean square error (RMSE), described by:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}, \quad (6)$$

where  $n$  denotes the number of sampled points in a sequence, and  $\hat{y}$  and  $y$  are the reconstructed output and the actual, respectively. The VAE indicator is constructed, therefore, by measuring the reconstruction error according to the RMSE, defined as the variational autoencoder reconstruction error (VAERE).

### 3.2 Principal component analysis

Principal Component Analysis (PCA) was first introduced in [44]. It is a methodology that obtains an  $r$ -dimensional basis that best captures the variance in the data. Given input data  $\mathbf{D} \in \mathbb{R}^{n \times d}$  and the desired threshold  $\alpha$ , it selects the smallest set of dimension  $r$  that captures at least an  $\alpha$  fraction of the total variance. The steps are shown in Algorithm 1.

The trajectory with the largest projected variance is called the first principal component, the orthogonal trajectory to the first one which captures the second largest projected variance is the second principal component, and so on.

We select a number of dimensions which is fewer than the one in the original dataset such as the subspace extent of these  $r$  dimensions captures at least  $\alpha$  fraction of the variance. In the practice,  $\alpha$  is set to a number around 0.95 (as it is in this case), so that the reduced dataset captures at least 95% of the total variance. Also,  $r=1$  so the original three time series are reduced to one that captures the features of the three with the minimum possible error.

## 4 Results

We now apply the methods in Section 3 to the observed sensor data. For the purpose of this study we perform univariate analysis only with SAT, then we perform multivariate analysis with all variables controlling SAT by reducing the dimensionality before applying these techniques.

Although these techniques have not been implemented in similar BMS-related failures so far, the techniques proved successful for early failure detection/tipping point analysis in similar dynamical systems such as manufacturing, automotive industry, climatological events, etc.

The different values for the windows can be chosen for each methodology according to the earliest they start to peak prior to failure. Once the values are chosen, the techniques can be extrapolated to similar problems (for instance early detection of failures in different parts of the HVAC system affecting room/area temperature).

### 4.1 Univariate analysis

Results for univariate analysis are presented in Figure 2, where SAT is presented together with the EWS indicators. The plot shows data from 64 hours prior to the failure at moment 0, which is when the temperature increases anomalously for the first time. We use different windows to obtain every EWS: 22 hours for variance and ACF1, 8 hours for PS and 4 hours for VAERE. The window size has been selected according to the clarity of the signal they provided on each indicator. Y-axis of the variance has been represented in logarithmic scale, for convenience.

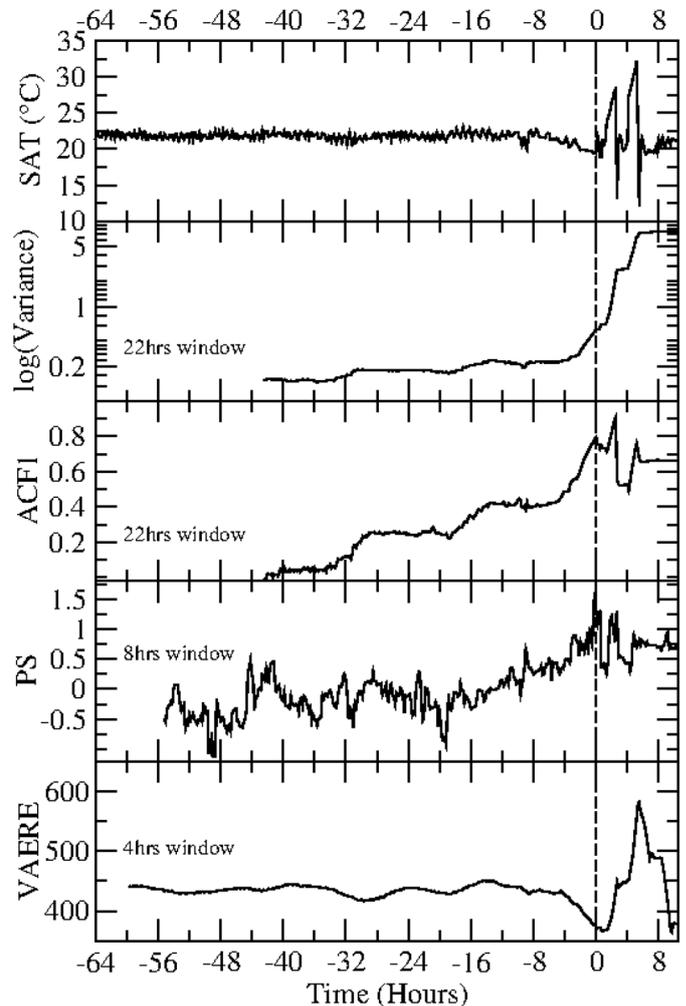
As shown in Figure 2, ACF1 and variance are the indicators presenting the most clear signals. It can be observed that PS also increases before the failure occurs, although it gives a signal with a shorter period. VAE has been first trained with the system functioning under normal conditions. The output displayed shows the VAERE. This seems to show a drop right before the tipping point, but less than 4 hours prior to event.

As can be seen in Figure 3, Kendalls show a more robust positive trend than negative, as the majority of values are greater than zero. However, some negative values can be seen as oscillations occur.

### 4.2 Multivariate analysis

Multivariate analysis for EWS is shown in Figure 4. The plot on the top corresponds to OAT, cooling valve and heating valve reduced with PCA to their first principal component or direction to the minimum projection variance. The windows used for each indicator for multivariate analysis are: 22 hours for variance and ACF1, 16 hours for PS and 20 hours for VAERE.

Comparing this plot with the univariate one, we observe that ACF1 and variance present clear EWS signals, whereas the variance represents a more steady slope in the multivariate case. PS shows a more prompt signal in the multivariate case, 24 hours prior to failure in comparison to 8 hours prior to the failure in the univariate

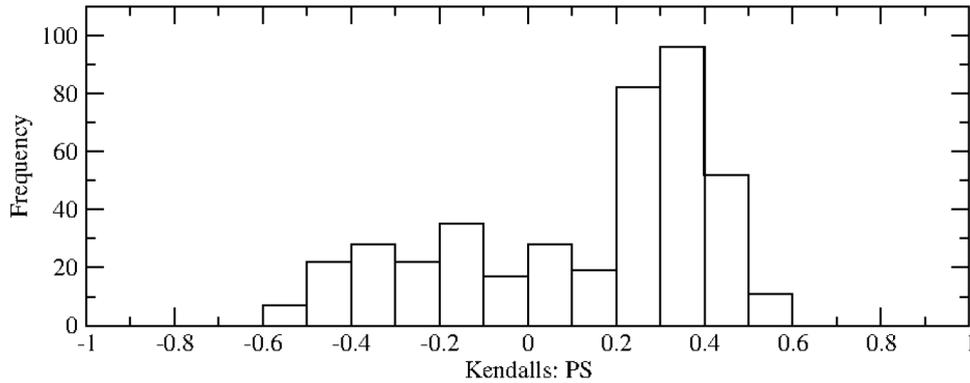


**Fig. 2.** Supply air temperature according to four indicators: variance, auto-correlation function, power spectrum and variational autoencoder. The moment of the failure (when the supply air temperature increases anomalously for the first time) is 9:30AM of the 27th of January 2019. Plot starts at 5:30 AM of the 25th of January, 64 hours prior to the failure. Variance is shown in logarithmic scale to make the increase prior to failure more visible. Window sizes used: 22 hours for variance and ACF1, 8 hours for PS and 4 hours for VAERE.

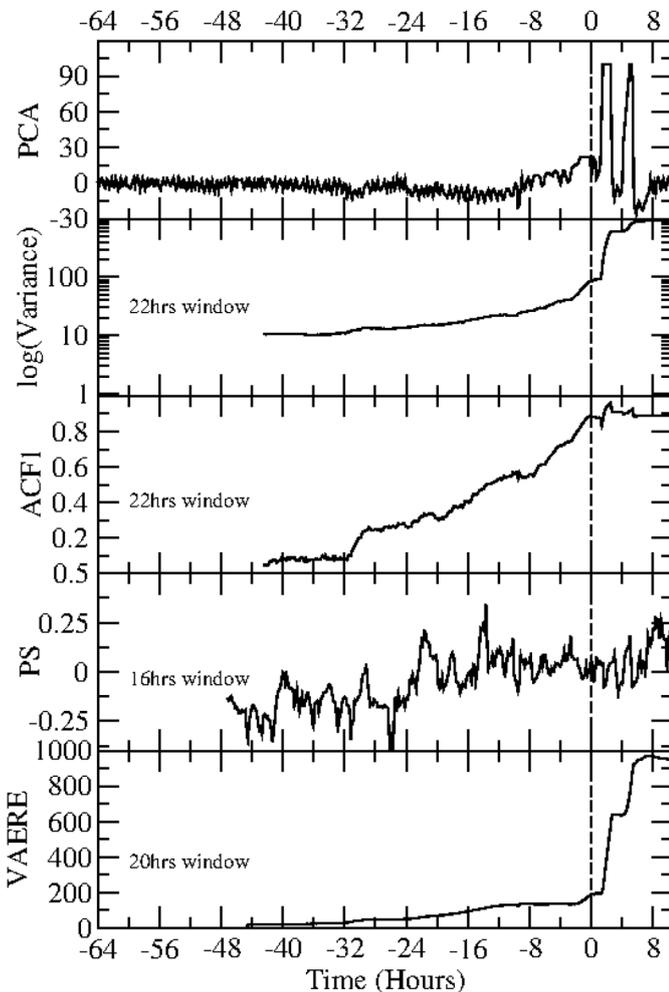
analysis. However, values of the Kendalls incline a bit more towards negative values than in the univariate case, although still positive values are predominant. This means that PS indicator is less robust in the multivariate case. An earlier signal is also given by the VAE, whose indicator for the multivariate signal gives a clear EWS after 32 hours before the event, in comparison to the 4 hours prior to failure of the univariate case.

## 5 Discussion, conclusions and limitations

In this paper, we applied several early warning signals techniques to analyse BMS temperature sensor data and percentage of openness of heating and cooling valves to



**Fig. 3.** Histogram of the frequency distribution of the Kendall trend statistic for the PS indicator for the univariate case, for the different windows across the indicator.

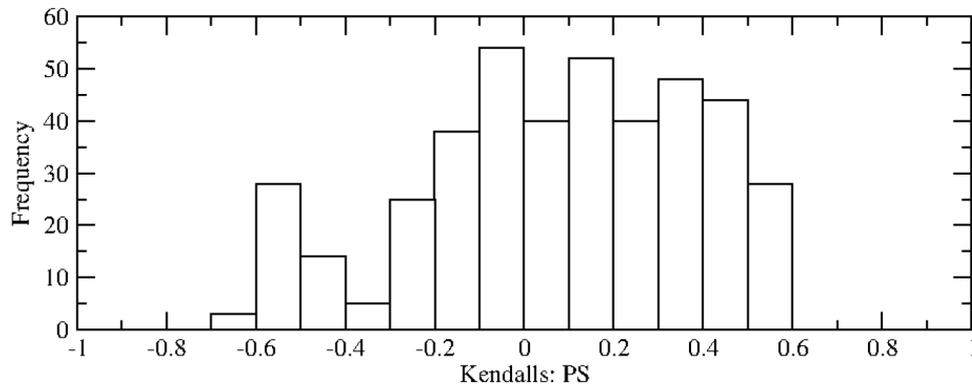


**Fig. 4.** Multivariate EWS analysis of the temperature failure with PCA and four indicators: variance, auto-correlation function, power spectrum and variational autoencoder. The moment of the failure (when the supply air temperature increases anomalously for the first time) is 9:30 AM of the 27th of January 2019. Plot starts at 5:30AM of the 25th of January, 64 hours prior to the failure. Variance is shown in logarithmic scale to make the increase prior to failure more visible. Window sizes used: 22 hours for variance and ACFI, 16 hours for PS and 20 hours for VAERE.

activate heating and/or cooling systems, respectively, in an area of a large industrial facility. The studied failure is a malfunction of one specific AHU control system that causes temperature spikes of up to 30 degrees celsius due to overreaction of the heating and cooling valves in response to an anomalous temperature change caused by the pre-heat coil in winter period in a specific manufacturing facility. Therefore the aim of the paper is to detect this anomaly before it happens. The analysis has been divided into univariate and multivariate EWS analysis, using for each the same indicators: variance, ACF, PS and VAE. PCA has been used for dimensionality reduction in the multivariate case.

The analysis shows that, in general, the indicators provide an earlier and more reliable signal in the multivariate case. This improvement can be seen specially in PS and VAE. In the case of PS, the difference is that the indicator completely changes its range of movement 24 hours prior to failure. With the VAE, the reconstruction error does not reproduce the gradual increase before the failure in the univariate case despite of the abrupt temperature change. In the multivariate case, the VAE does show the expected behaviour, producing an early signal almost 32 hours prior to failure. The reason that the VAE does not provide a good signal in the univariate case may lay in what is considered a normal functioning system when training the model using only SAT. Such distribution makes small changes due to external factors, which actually control its non-stationary behaviour. In the multivariate case, the behaviour of the system is defined mainly by the heating valve, which steadily opens and closes to provided heating to the environment within an established range or set point. In both univariate and multivariate analysis, ACF demonstrated to perform very well, starting to clearly go above its 0.2 value almost 32 hours prior to failure, with a very similar performance in both cases.

The SAT is controlled by average of the four temperature sensors in the room. The choice of the SAT instead of each the four temperature sensors is to avoid false positives due to noise. Although there is a standard on the locations on which the sensors should be installed, they can be temporarily exposed to heat sources (if someone places a desk with a computer next to it by mistake or



**Fig. 5.** Histogram of the frequency distribution of the Kendall trend statistic for the PS indicator for the multivariate case, for the different windows across the indicator.

sunlight directly beaming at the sensor, for instance). Therefore there could have been external factors that could not have been controlled for the purpose of this study.

The indicators generally respond earlier in the multivariate analysis, as the variables controlling SAT are the ones considered for the dimensionality reduction prior to applying the EWS techniques. This has been possible due to a good knowledge of the system so the right variables controlling SAT, out of hundreds of sensors, could be selected. In other circumstances the lack of knowledge about the system or sensors installed would not create the conditions for multivariate analysis. In such case, a technique for single variable EWS analysis can also deliver satisfactory results.

In this work, lack of data and difficulty of automation are the main limitations encountered. Here the problem lies not only in the automation of the data acquisition production, but also in the location of specific failures in order to test the algorithms. Further work to improve these results would imply an improvement of software solutions for data extraction, as well as the location of new failures and case studies to further improve and generalise these results.

The early detection of such failures gives time to on-site engineers to make adjustments when necessary before these failures actually happen. This not only reduces maintenance and operational costs, but also produces energy savings by advising when parts of the system should not be activated in some given period, thus compensating the “blind spots” from the BMS control system by extracting real value from the data generated.

## Funding information

This work was funded by EPSRC DTP PhD Studentship, NPL and Mitie. For the purpose of Open Access the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Nomenclature

ACF	Auto-Correlation Function
AHU	Air Handling Unit

ANN	Artificial Neural-Network
APAR	Air handling unit Performance Assessment Rules
BMS	Building Management Systems
DFA	Detrended Fluctuation Analysis
EOF	Empirical Orthogonal Functions
EWS	Early Warning Signals
FDD	Fault Detection and Diagnosis
GRNN	General Regression Neural-Network
HVAC	Heating, Cooling and Air Conditioning
OAT	Outside Air Temperature
PCA	Principal Component Analysis
PS	Power Spectrum
SAT	Supply Air Temperature
VAE	Variational Auto Encoder
VAERE	Variational Auto Encoder Reconstruction Error

## References

1. L. Alessi, C. Detken, ‘Real time’ early warning indicators for costly asset price boom/bust cycles: a role for global liquidity, 2009
2. D. Gerdesmeier, H.-E. Reimers, B. Roa et al., Early warning indicators for asset price booms, *Rev. Econ. Finance* **3**, 1–20 (2011)
3. J. Wen, H. Gao, Degradation assessment for the ball screw with variational autoencoder and kernel density estimation, *Adv. Mech. Eng.* **10**, 1687814018797261 (2018)
4. V.N. Livina, T.M. Lenton, A modified method for detecting incipient bifurcations in a dynamical system, *Geophys. Res. Lett.* **34** (2007)
5. M. Dey, S.P. Rana, S. Dudley, A case study based approach for remote fault detection using multi-level machine learning in a smart building, *Smart Cities* **3**, 401–419 (2020)
6. V. Livina, A. Lewis, M. Wickham, Tipping point analysis of electrical resistance data with early warning signals of failure for predictive maintenance, [arXiv:1904.04636](https://arxiv.org/abs/1904.04636) (2019)
7. D.P. Kingma, M. Welling, Auto-encoding variational bayes, [arXiv:1312.6114](https://arxiv.org/abs/1312.6114) (2013)
8. K. Pearson, Principal components analysis, *London Edinburgh Dublin Philos. Mag. J. Sci.* **6**, 559 (1901)
9. J. Schein, S.T. Bushby, N.S. Castro, J.M. House, A rule-based fault detection method for air handling units, *Energy Build.* **38**, 1485–1492 (2006)

10. P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: an explanation of the  $1/f$  noise, *Phys. Rev. Lett.* **59**, 381 (1987)
11. F. Titikpina, A. Caucheteux, A. Charki, D. Bigaud, Uncertainty assessment in building energy performance with a simplified model, *Int. J. Metrol. Quality Eng.* **6**, 308 (2015)
12. M.G. Kendall, Rank correlation methods (1948)
13. D.P. Kingma, M. Welling et al., An introduction to variational autoencoders, *Found. Trends Mach. Learn.* **12**, 307–392 (2019)
14. J.M. Drake, B.D. Grien, Early warning signals of extinction in deteriorating environments, *Nature* **467**, 456–459 (2010)
15. B.M. Rogers, K. Solvik, E.H. Hogg, J. Ju, J.G. Masek, M. Michaelian, L.T. Berner, S.J. Goetz, Detecting early warning signals of tree mortality in boreal north America using multiscale satellite data, *Glob. Change Biol.* **24**, 2284–2304 (2018)
16. J. Broderick, R. Burchill, H. Clark, A system for early warning of bearing failure (1972)
17. B. Schwaab, S.J. Koopman, A. Lucas, Systemic risk diagnostics: coincident indicators and early warning signals (2011)
18. G.E. Box, G.M. Jenkins, G.C. Reinsel, G.M. Ljung, Time series analysis: forecasting and control, John Wiley & Sons (2015)
19. Y. Zhou, C. Lork, W.-T. Li, C. Yuen, Y.M. Keow, Benchmarking air-conditioning energy performance of residential rooms based on regression and clustering techniques, *Appl. Energy* **253**, 113548 (2019)
20. V.N. Livina, F. Kwasiok, T.M. Lenton, Potential analysis reveals changing number of climate states during the last 60 kyr, *Climate Past* **6** (2010)
21. J.T. Andrews, E.J. Morton, L.D. Grin, Detecting anomalous data using auto-encoders, *Int. J. Mach. Learn. Comput.* **6**, 21 (2016)
22. D. Bigaud, A. Charki, A. Caucheteux, F. Titikpina, T. Tiplica, Detection of faults and drifts in the energy performance of a building using Bayesian networks, *J. Dyn. Syst. Measur. Control* **141** (2019)
23. W.-Y. Lee, J.M. House, N.-H. Kyong, Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks, *Appl. Energy* **77**, 153–170 (2004)
24. Y. Qin, C. Yuen, S. Adams, Invariant learning based multi-stage identification for lithium-ion battery performance degradation, [arXiv:2008.05123](https://arxiv.org/abs/2008.05123) (2020)
25. L. Alessi, C. Detken, Quasi real time early warning indicators for costly asset price boom/bust cycles: a role for global liquidity, *Eur. J. Political Econ.* **27**, 520–533 (2011)
26. C. Matt, T. Hess, A. Benlian, Digital transformation strategies, *Bus. Inf. Syst. Eng.* **57**, 339–343 (2015)
27. D. Paschek, C.T. Luminosu, A. Draghici, Automated business process management in times of digital transformation using machine learning or artificial intelligence, in: MATEC Web of Conferences, Vol. 121, EDP Sciences, 2017, p. 04007
28. S. Yang, M.P. Wan, W. Chen, B.F. Ng, S. Dubey, Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization, *Appl. Energy* **271** (2020)
29. J. Qiu, W. Liang, L. Zhang, X. Yu, M. Zhang, The early-warning model of equipment chain in gas pipeline based on DNN-hmm, *J. Natur. Gas Sci. Eng.* **27**, 1710–1722 (2015)
30. V. Livina, G. Lohmann, M. Mudelsee, T.M. Lenton, Forecasting the underlying potential governing the time series of a dynamical system, *Physica A* **392**, 3891–3902 (2013)
31. H. Stopps, M. Touchie, Reduction of HVAC system runtime due to occupancy-controlled smart thermostats in contemporary multi-unit residential building suites, in IOP Conference Series: Materials Science and Engineering, Vol. 609, IOP Publishing, 2019
32. W.-T. Li, N.U. Hassan, F. Khan, C. Yuen, Y.M. Keow, Data driven model for performance evaluation and anomaly detection in integrated air source heat pump operation, in ICIT, 2019, pp. 1280–1285
33. A. Ari, R. Dagtekin, Early warning signals of the 2000/2001 Turkish Nancial crisis (2007)
34. Y.-C. Lu, C.-H. Shen, Y.-C. Wei, Revisiting early warning signals of corporate credit default using linguistic analysis, *Pacific-Basin Finance J.* **24**, 1–21 (2013)
35. E.P. Davis, D. Karim, Comparing early warning systems for banking crises, *J. Financial Stab.* **4**, 89–120 (2008)
36. U.B. Singh, A. Ahluwalia, C. Sharma, R. Jindal, R. Thakur, Planktonic indicators: a promising tool for monitoring water quality (early-warning signals), *Ecol. Environ. Conserv.* **19**, 793–800 (2013)
37. T. Squartini, I. Van Lelyveld, D. Garlaschelli, Early-warning signals of topological collapse in interbank networks, *Sci. Rep.* **3**, 3357 (2013)
38. W.-Y. Lee, J.M. House, C. Park, G.E. Kelly, Fault diagnosis of an air-handling unit using artificial neural networks, *Trans. Am. Soc. Heating Refriger. Air Condition. Eng.* **102**, 540–549 (1996)
39. J. Prettyman, T. Kuna, V. Livina, A novel scaling indicator of early warning signals helps anticipate tropical cyclones, *Europhys. Lett.* **121**, 10002 (2018)
40. M. Dey, S.P. Rana, S. Dudley, Smart building creation in large scale hvac environments through automated fault detection and diagnosis, *Fut. Gener. Comput. Syst.* **108**, 950–966 (2020)
41. Y. Yu, D. Woradechjumbo, D. Yu, A review of fault detection and diagnosis methodologies on air-handling units, *Energy Build.* **82**, 550–562 (2014)
42. T. Lenton, V. Livina, V. Dakos, E. Van Nes, M. Scheer, Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness, *Philos. Trans. Royal Soc. A* **370**, 1185–1204 (2012)
43. J. Prettyman, T. Kuna, V. Livina, Generalized early warning signals in multivariate and gridded data with an application to tropical cyclones, *Chaos* **29**, 073105 (2019)
44. M. Li, T. Zeng, R. Liu, L. Chen, Detecting tissue-specific early warning signals for complex diseases based on dynamical network biomarkers: study of type 2 diabetes by cross-tissue analysis, *Briings Bioinform.* **15**, 229–243 (2014)
45. M.J. Zaki, W. Meira, Data mining and analysis: fundamental concepts and algorithms, Cambridge University Press (2014)