

## Adverse Condition and Critical Event Prediction in Commercial Buildings: Danish Case Study

Egedorf, Søren; Shaker, Hamid Reza; Martin, Rodney A. ; Jørgensen, Bo Nørregaard

*Published in:*  
Energy Informatics

*DOI:*  
[10.1186/s42162-018-0015-5](https://doi.org/10.1186/s42162-018-0015-5)

*Publication date:*  
2018

*Document version:*  
Final published version

*Document license:*  
CC BY

*Citation for pulished version (APA):*  
Egedorf, S., Shaker, H. R., Martin, R. A., & Jørgensen, B. N. (2018). Adverse Condition and Critical Event Prediction in Commercial Buildings: Danish Case Study. *Energy Informatics*, 1(1), Article 10.  
<https://doi.org/10.1186/s42162-018-0015-5>

Go to publication entry in University of Southern Denmark's Research Portal

### Terms of use

This work is brought to you by the University of Southern Denmark.  
Unless otherwise specified it has been shared according to the terms for self-archiving.  
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.  
Please direct all enquiries to [puresupport@bib.sdu.dk](mailto:puresupport@bib.sdu.dk)

RESEARCH

Open Access



# Adverse condition and critical event prediction in commercial buildings: Danish case study

Søren Egedorf<sup>1</sup>, Hamid Reza Shaker<sup>1\*</sup> , Rodney A. Martin<sup>2</sup> and Bo Nørregaard Jørgensen<sup>1</sup>

\*Correspondence:

[hrsh@mmmi.sdu.dk](mailto:hrsh@mmmi.sdu.dk)

<sup>1</sup>Center for Energy Informatics,  
Mærsk Mc–Kinney Møller Institute,  
University of Southern Denmark,  
Campusvej, 5230 Odense M,  
Denmark

Full list of author information is  
available at the end of the article

## Abstract

Over the last two decades, there has been a growing realization that the actual energy performances of many buildings fail to meet the original intent of building design. Faults in systems and equipment, incorrectly configured control systems and inappropriate operating procedures increase the energy consumption about 20% and therefore compromise the building energy performance. To improve the energy performance of buildings and to prevent occupant discomfort, adverse condition and critical event prediction plays an important role. The Adverse Condition and Critical Event Prediction Toolbox (ACCEPT) is a generic framework to compare and contrast methods that enable prediction of an adverse event, with low false alarm and missed detection rates. In this paper, ACCEPT is used for fault detection and prediction in a real building at the University of Southern Denmark. To make fault detection and prediction possible, machine learning methods such as Kernel Density Estimation (KDE), and Principal Component Analysis (PCA) are used. A new PCA-based method is developed for artificial fault generation. While the proposed method finds applications in different areas, it has been used primarily for analysis purposes in this work. The results are evaluated, discussed and compared with results from Canonical Variate Analysis (CVA) with KDE. The results show that ACCEPT is more powerful than CVA with KDE which is known to be one of the best multivariate data-driven techniques in particular, under dynamically changing operational conditions.

**Keywords:** Building energy performance, Adverse condition and critical event prediction, Artificial fault generation, Fault detection and prediction, Canonical variate analysis, Machine learning

## Introduction

Over the last decade, the contribution of buildings energy consumption to total energy consumption has been between 20% – 40% in developed countries (Lombard et al. 2007; Shaker and Lazarova-Molnar 2017). Today the figure points towards a contribution of around 40%. In addition to this, buildings account for approximately 20% of total CO<sub>2</sub> emissions (Lazarova-Molnar et al. 2016). Thus, there is an excellent opportunity for reducing energy consumption and CO<sub>2</sub> emissions if the general performance of energy-consuming equipment in buildings could be improved.

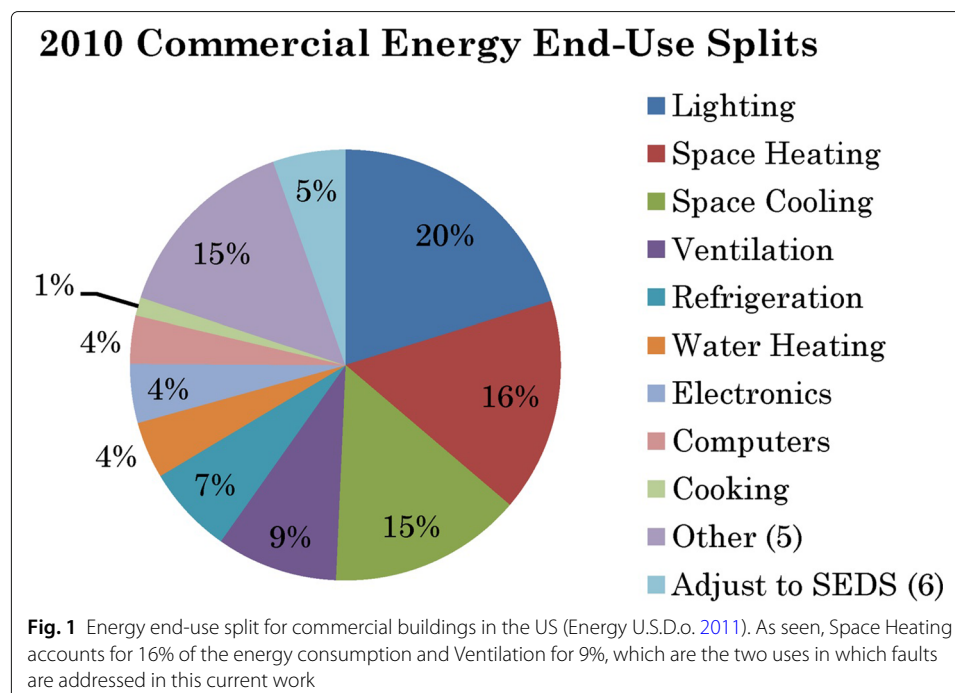
A traditional, and more passive measure for improving energy performance of buildings is to implement energy conservation measures such as more insulation to exterior

walls, ceilings and floors, new insulating windows etc. (Tommerup et al. 2004) which is important. However, with the emergence of new and smarter buildings and new intelligent building equipment, new measures could be implemented. The Danish government aims at a reduction in energy consumption in new buildings in 2020 by 75% relative to 2006 levels. In addition, by 2050 the energy consumption should be reduced by 50% in existing buildings (Government 2009). Thus, there is room for new and innovative solutions for reducing energy consumption to reach these goals (Jørgensen et al. 2015). Faults in buildings compromise the energy performance and also cause occupants discomfort. There are different faults in buildings. Examples are duct leakages in ventilation system, simultaneous heating/cooling, and dampers in ventilation system not working properly (Lazarova-Molnar et al. 2016). Thus, there is a need to detect those faults early so their impact on energy consumption will be minimized.

In the U.S., the total energy consumption in commercial buildings has been divided into different end-uses as shown in Fig. 1<sup>1</sup>. This figure shows how expensive a fault can be in terms of its energy use. Furthermore, studies show that 25% – 45% of HVAC energy consumption is wasted due to faults (Akinci et al. 2011) and the most typical faults in commercial buildings are the ones shown in Fig. 2 (Roth et al. 2005), which is a table of the annual impact of each fault in terms of energy consumption.

Studies have shown that in 2009, only 13 of the most common faults in buildings have caused over \$3.3 billion in energy waste in the U.S. (Mills 2011).

To improve the energy performance of the buildings, Fault Detection and Diagnosis (FDD) methods are used. However, fault detection and diagnosis in buildings are challenging tasks. Early detection of faults and adverse conditions has been the subject of research in many fields. NASA Ames recently has released a tool which is called ACCEPT which has shown to be very effective at comparing methods used for fault detection and prediction. ACCEPT has shown a good performance compared to other state-of-the-art



Fault	Fault type	By percentage	Annual energy consumption [quads]
Duct leakage	Air Distribution	30%	0.3
HVAC left on when space unoccupied	HVAC	20%	0.2
Lights left on when space unoccupied	Lighting	18%	0.18
Airflow not balanced	Air Distribution	7%	0.07
Improper refrigerant change	Refrigeration Circuits	7%	0.07
Dampers not working properly	Air Distribution	6%	0.055
Insufficient evaporator airflow	Air Distribution	4%	0.035
Improper controls setup/commissioning	Controls	2%	0.023
Control component failure or degradation	Controls	2%	0.023
Software programing error	Controls	1%	0.012
Improper controls hardware installation	Controls	1%	0.010
Air-cooled condenser fouling	Refrigeration Circuits	1%	0.008
Valve leakage	Waterside Issues	1%	0.007
Total	-	100%	1.0

**Fig. 2** The annual impact of faults in terms of energy consumption (Roth et al. 2005)

methods (Egedorf and Shaker 2017), when applied to data from Cranfield Multiphase Flow Facility (Ruiz-Cárcel et al. 2015) which is a well-known benchmark example.

In this paper, ACCEPT is used for fault detection and prediction in buildings. A combined office and classroom building at SDU is used to evaluate the performance of ACCEPT in detecting and predicting faults. The performance of a method is determined by its False Alarm Rate (FAR), Missed Detection Rate (MDR) and Detection Time (DT) which are explained briefly in the paper. In order to allow for the data from the building to be used in ACCEPT, methods such as KDE and PCA-based contribution plots are used. A new PCA-based method is also developed and introduced for artificial fault generation. While the proposed method finds applications in different areas, it has been used for analysis purposes in this work. The results from ACCEPT are evaluated, discussed and compared with results from CVA with KDE. CVA with KDE has proven to be one of the best performing state-of-the-art methods in FDD - both on the Tennessee Eastman Process Plant (Odiowei and Cao 2010) and on Cranfield Multiphase Flow Facility (forming a real world data set) (Ruiz-Cárcel et al. 2015) - and thus forms a good basis for comparison.

### Description of case study: building OU44 in University of Southern Denmark (SDU)

The case study which has been selected is building OU44 from SDU. The data extracted includes six different physical measures in each of four different rooms from the building. Thus, the data contains  $4 \times 6 = 24$  different variables and the data are captured from beginning of September 2016 to end of January 2017. The four rooms are the following:

- Ø20-601b-2 (Classroom on 2. floor).
- Ø20-604b-1 (Study room on 1. floor below Ø20-601b-2).
- Ø22-508-1 (Study room on 1. floor).
- Ø22-511-2 (Classroom on 2. floor above Ø22-508-1).

Refer to the SDU webpage<sup>2</sup> for a drawing of the building with the rooms. The physical measures from each room are:

- CO<sub>2</sub>: CO<sub>2</sub>level in the room measured in ppm.

- Radiator\_valve: Degree of opening of heating unit.
- Temperature: Temperature in the room measured in °C.
- Valve\_control\_from\_CO2: Desired degree of opening of ventilation unit due to CO<sub>2</sub>, (increase ventilation if CO<sub>2</sub> ppm is too high).
- Valve\_control\_from\_temperature: Desired degree of opening of ventilation unit due to temperature, (increase ventilation if temperature is too high).
- Valve: Degree of opening of ventilation unit.

The 24 variables are named and numbered according to the above two lists.

### Data preprocessing

The data capturing frequency is threshold based, which means that new measurements are only captured when a measure reaches different thresholds. For temperatures, this threshold is usually 0.1 °C, and thus if temperature is constant for several hours, no data is captured, but when it increases/decreases by 0.1 °C a new measurement is captured. Thus, the data needs to be re-sampled to a common frequency, to make the measures from the 24 variables correspond to the same time instances. The common frequency used is 5 min intervals, which means that 25921 observations are presented in the data set, spanning three months. Because of this re-sampling, some of the variables are constant for longer periods. The data has therefore been linearly interpolated. Since it becomes clear that variables 17, 18 and 19 do not have any information, they have been removed and we now have 21 variables.

Furthermore, the preprocessed data is used for training since no faults are present. However, a testing and also a validation data set are required by ACCEPT. CVA requires only a testing data set. Thus, a set of faulty data needs to be generated. We have done this through introducing artificial faults. This method will be further elaborated on in the methodology section. Of course, real data could be used as well by deliberately seeding faults to the building. However, this could potentially cause reduced comfort followed by complaints by occupants in addition to the time consuming process of collecting the faulty data. Before selecting/developing an Artificial Fault Generation (AFG) method, a review on how to artificially generate data is presented.

### On artificial faults generation

One of the prevailing scientific paradigms in data-driven FDD research is to develop better methods (or improve on existing methods) and test their performance against known data sets. One of the most widely used data sets are those from the Tennessee Eastman process. Thus, data for training and testing purposes does exist to validate and compare different methods. Other approaches to generate data include model based computer simulation (Kothamasu et al. 2004; Rodríguez et al. 2008), data acquired in small test rigs or particular parts of a physical system (Ruiz-Cárcel et al. 2015). However, what to do when data from a system is insufficient or missing - that is, how to handle cases in which we face the challenge of lack of data and in this case faulty data? Of course one could carry-out physical tests, as in (Ruiz-Cárcel et al. 2015), to generate faulty data, but that would require a considerable amount of work or might not be possible due to restrictions (as mentioned; complaints by occupants using the building). Another approach could be to develop a mathematical model and perform simulations, but the complexities involved

in simulating the real system are prohibitive. Since training data is available, as mentioned, the problem seems to be to generate artificial testing and validation data sets from the training data set.

A method of the generation of artificial data is known as Virtual Sample Generation (VSG). The key problem this method tries to solve is the small data-set learning problem. That is, when the training data sample sizes are small, a biased learning results will be obtained. VSG can help to avoid this. Recent research on VSG is presented in (Li et al. 2017) and (Sha et al. 2013), but to be applicable to this current work, modifications are needed. The main issue to address is that VSG generates virtual data based on knowledge from a small data set, and the generated larger virtual data set then shares the same distribution (or Membership Function as in (Li et al. 2017)) as the small data set. We need the method to generate a faulty data set that exhibit different distributions (in faulty regions) than the healthy training data set, as well as the method needs to take into account the correlations in the data. As such, we want to address the following questions: How much would the CO<sub>2</sub> level change if the temperature drops by 3 °C? and how would the rest of the variables change? Considering the complexity of this, modified VSG will not be used in this current work, but another method is developed to address the mentioned needs.

The last remaining question would be - at what time instances and how often do the faults occur? For this work, a simple approach is taken, but for future work a more complex approach was identified in the literature, such as the method in (Zhang et al. 2015), known as Fault Sample Generation (FSG). This is the study of how faults occur randomly, which follow certain statistical distributions and properties (such as the average lifetime of components). However, because of the complexity of these methods and scope of this work a simple approach is taken - faults occur on a random weekday every week for 12 weeks. Three typical faults will be considered. These faults are: an open window during night, an open window during day and finally a ventilation fault during day. More details about artificial generation of these faults will be explained later in the section describing the results.

## Methods

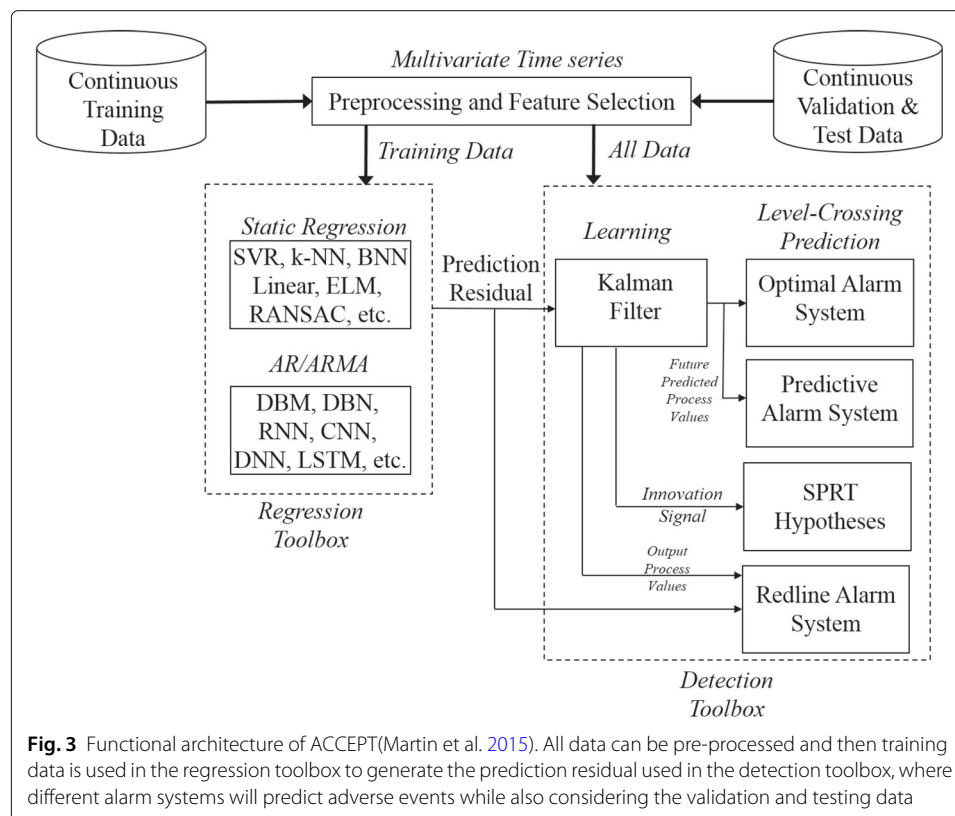
The methods used include ACCEPT, as documented in (Martin et al. 2015) as well as KDE, as documented in (Ruiz-Cárcel et al. 2015) and (Odiwei and Cao 2009). PCA-based contribution plots will also be used to determine the variables mostly contributing to the faults (since ACCEPT requires a variable to predict). Then KDE is used to estimate the probability density functions of the relevant variables (depending on what the contribution plot shows) in the training data sets to develop an empirical value for the ground truth. Finally, a PCA-based AFG method is presented.

### A brief description of ACCEPT

In short, ACCEPT is a generic MATLAB-Based framework for adverse effect and critical event prediction. ACCEPT is an architectural framework developed to compare and contrast the performance of a variety of machine learning and early warning algorithms. ACCEPT tests and compares these algorithms according to their ability to predict adverse events in arbitrary time-series data from systems or processes. This ability (or performance) is measured using previously mentioned metrics such as MDR, FAR and DT (Martin et al. 2015).

ACCEPT is patterned after, and shares the same basic composition as the Multivariate State Estimation Technique (MSET). MSET is an existing state-of-the-art method used for prediction of adverse events in advance of their occurrence and was originally used in nuclear applications and aviation and space applications. However, as distinct from MSET, ACCEPT is an *open-source* tool that offers users to choose from a variety of machine learning algorithms that can be tuned via hyperparameter optimization using the regression toolbox. Furthermore, additional detection algorithms based upon hypothesis testing go beyond the standard SPRT (Sequential Probability Ratio Test) hypotheses offered by MSET in the detection toolbox.

As shown in Fig. 3, all data can be pre-processed which basically means that each variable in the multivariate data will be centered to zero mean and scaled to unit variance using z-score normalization. Normalizing the multivariate data can be important since the data consists of different variables (or features) and each variable has a different physical meaning. Feature selection is the process of selecting only the variables relevant to the process being monitored - some variables may have no relevant information and these should be removed before performing operations on the data (Chiang et al. 2001). In doing so it will reduce computational burden, make models easier to interpret by simplification, reduce overfitting and avoiding the curse of dimensionality (Bolón-Canedo et al. 2015; Tuv 2009; Okun 2011). Usually, feature selection is not necessary with low dimensions as in this work with only 21 features. It was found in this work that feature selection and normalizing data was not necessary as the satisfying results from ACCEPT was achieved without feature selection





The regression toolbox, represented on the left of the figure, contains many regression algorithms from which to generate the output of this box - the prediction residual based on training data. The chosen algorithm (by the user) processes a number of features - the multivariate time-series - and predicts a chosen target parameter based on these input features and compares this prediction with the actual value to generate the prediction residual. This mapping of the target parameter characterizes the basic relationship/correlation between the input features and the target parameter (or response variable) for regression. Thus, it is important that the input features are adequate predictors of the target variable (Martin et al. 2015).

As mentioned, the prediction residual quantifies the difference between the actual value of the target parameter and the predicted value. An optimization problem is established and this problem is essentially the result of a so-called  $f$ -fold cross validation. The Normalized Mean Squared Error (NMSE) is the objective function of this optimization problem subject to a regression specific hyper parameter. The NMSE of resulting residuals represents regression performance and is minimized when generating the residual (Martin et al. 2015). The lower the NMSE, the better the regression performance, although one needs to take care to prevent over-fitting by acknowledging bias-variance tradeoff and “detuning” when necessary.

In the detection toolbox (or step), a validation data set containing occurrences of adverse events is used in the design of an alarm system. This data set should in theory be drawn from the same distribution as the final testing data set which also contains adverse events (Brutsaert et al. 2016). All detection algorithms will use Receiver Operating Characteristics curves (ROC curve) analysis to enable the design of trade-offs between FAR and MDR, and in all cases an equal trade-off will be used. All detection methods used are threshold based and thus, if the resulting threshold from the ROC curve analysis is crossed, an alarm is triggered. The performance metrics that ACCEPT produces are defined as follows:

- FAR - An alarm is triggered at a time point that does not contain an example of a confirmed anomalous event in at least one time point in the next  $d$  time steps (Martin et al. 2015).
- MDR - No alarm is triggered at a time point where an example of a confirmed anomalous event exists in at least one time point in the next  $d$  time steps (Martin et al. 2015).
- DT - Time steps prior to the occurrence of a future adverse event, which is detected by the prediction system (Brutsaert et al. 2016).

Prior to using the ROC curve for design purposes, an optimization problem is established to maximize the Area Under the ROC Curve (AUC). A Linear Dynamical System (labeled as “Kalman Filter”) is obtained from the residual output, and both the learned LDS parameters derived from training data and the adverse events contained in the validation data set are used in the optimization. The AUC optimization problem is parameterized by the state dimension of the LDS  $n$  and the prediction horizon  $d$ , taking values of  $n_{opt} = 2$  and  $d_{opt} = 1$ , respectively.

Note that the AUC optimization problem is only the first step in determining the threshold and is conducted to find the LDS state dimension  $n$  and prediction horizon  $d$  that produces the highest AUC value. The next step is to use the produced ROC curve for



selecting the threshold, and as mentioned an equal trade-off between MDR and FAR will be used for design purposes in all cases. The threshold selected is ultimately the goal of producing the most accurate representation of the ground truth (Martin et al. 2015). The following regression techniques will be studied in this work:

- Linear Ridge Regression (LIN)
- Extreme Learning Machine (ELM)

and the following detection algorithms:

- Redline - Training (RT)
- Redline - Validation (RV)
- Predictive - Training (PT)
- Predictive - Validation (PV)
- Optimal - Training (OT)
- Optimal - Validation (OV)

#### Description of kernel density estimation

The probability of a random variable  $x$  (with a probability density function  $p(x)$ ) to be smaller than a certain value  $s$  is defined as:

$$P(x < s) = \int_{-\infty}^s p(x) dx \quad (1)$$

This equation is used to determine the ground truth limit, for a target variable, by solving the equation  $P(x < s) = 1 - \alpha/2$ , where  $\alpha$  is the significance level and  $s$  is the solution. This means that the value  $s$  determines that  $(1 - \alpha/2)100\%$  of the data lies at a lower value than  $s$ . In the case where the lower limit should be used in the ground truth function,  $P(x < s) = \alpha/2$  is solved. Here,  $p(x)$  can be calculated through the kernel function  $K$ :

$$p(x) = \frac{1}{Mh} \sum_{k=1}^M K\left(\frac{x - x_k}{h}\right) \quad (2)$$

$$K(g) = e^{-g^2/2} / \sqrt{2\pi} \quad (3)$$

where  $h$  is the selected bandwidth (see (Odiwei and Cao 2009)),  $M$  is the sample size and  $x_k$  is the  $k^{\text{th}}$  sample of  $x$ . By replacing  $x_k$  with the sample variable of interest, it is possible to estimate the probability density function of this variable (Ruiz-Cárcel et al. 2015). There is no single way of selecting a correct  $h$  for a given application, but it is important to ensure that the estimated distribution is not too rough or too flat which can be the case with a too small or too big  $h$  respectively (Odiwei and Cao 2009).

#### Determining contribution plots based on PCA

As mentioned, PCA-based contribution plots will be used to select the target parameter to be used by ACCEPT. PCA simplifies the monitoring of a process by converting the high-dimensional data, using loading vectors determined by a singular value decomposition, into lower-dimensional so-called score vectors which capture and preserve the spatial correlations between variables while also capturing most of the variation in the data. An elliptical confidence bound can be superimposed on the same plot containing the principal components. Retaining only the two first principal components is often sufficient to capture the most information from the data, thus making it possible to use a

two-dimensional Cartesian coordinate system. When the elliptical confidence bound is crossed a fault has been detected, and then the next step is to use contribution plots to determine the origin of the fault. That is, which variable is contributing mostly to the out-of-control status? Contribution plots are a PCA approach to fault identification, and it determines the contribution of each variable to the principal components determined by PCA. The contribution plot can be based on a single observation at a specific time instance, samples of observations, or on all data. The contribution of each variable  $x_j$  to the out-of-control scores  $t_i$  is calculated as

$$cont_{i,j} = \begin{cases} \frac{t_i p_{i,j}}{\sigma_i} \frac{x_j - \mu_j}{\sigma_j} & x_j \geq \mu_j \\ 0 & x_j < \mu_j \end{cases} \quad (4)$$

Where  $p_{i,j}$  is the  $(i,j)^{th}$  element of the loading matrix  $P$ ,  $\sigma_i$  is the corresponding singular value and  $\sigma_j$  and  $\mu_j$  is the standard deviation and mean of the variable  $x_j$ , respectively. The total contribution of the  $j^{th}$  process variable  $x_j$  is then calculated as (Chiang et al. 2001):

$$CONT_j = \sum_{i=1}^r (cont_{i,j}) \quad (5)$$

Where  $r$  is the number of score vectors or principal components retained. This  $CONT_j$  can then be plotted to illustrate the contributions of each variable to the fault. Like PCA-based contribution plots, CVA-based can also be used or combined with the PCA-based. However, it has been observed that the two plots usually shows the same variable contributing the most (Egedorf and Shaker 2017; Egedorf 2017). Therefore only the PCA-based is used in this work.

#### PCA-based Artificial Fault Generation (AFG)

In this section, a new PCA-based method is developed for artificial fault generation. While the proposed method finds applications in different areas, it has been used primarily for analysis purposes in this work. This method of introducing faults to the training data set is based on a PCA-method documented in (Chiang et al. 2001). Principal components are used to represent the healthy or faulty state of the system. The idea is to add faults to the components and then project these vectors back to the high dimensional space to get a faulty data set. Thus, the spatial correlations are preserved when adding faults to the training data set. As a first step the data is loaded in a matrix with  $m = 21$  process variables and  $n = 25921$  observations as shown in Eq. 6:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (6)$$

Then each of the 21 variables in the training data set are z-score normalized to 0 mean and standard deviation 1. A Singular Value Decomposition (SVD) is performed on the data as shown in Eq. 7

$$\frac{1}{\sqrt{n-1}} X = USV^T \quad (7)$$

Where  $U \in R^{n \times n}$  and  $V \in R^{m \times m}$  are unitary (orthogonal in this case) matrices and  $S \in R^{n \times m}$  contains the non-negative real singular values of decreasing magnitude along

its main diagonal ( $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq 0$ ). The loading vectors are the orthonormal column vectors in the matrix  $V$  and the variance of the training set projected along the  $i$ 'th column of  $V$  is equal to  $\sigma_i^2$ . Typically, the loading vectors corresponding to the  $a$  largest singular values are retained, where  $a$  can be determined by e.g. the *percent variance test* (Chiang et al. 2001). However, that is for process monitoring purposes, and in this case the purpose is to introduce artificial faults to the data set. Thus,  $a$  is set to 1, to add faults in only one vector. Therefore, selecting only the first  $a$  column vectors in  $V$  which captures the most of the variation in the data set, the loading matrix  $P \in R^{m \times a}$  can be formed. The projections of the observations in  $X$  into the lower-dimensional space are contained in the score matrix  $T$  which is formed as in Eq. 8:

$$T = XP \quad (8)$$

Where  $T \in R^{n \times a}$ . Projecting back to the  $m$  dimensional space yields:

$$\hat{X} = TP^T \quad (9)$$

The z-score normalization of  $\hat{X} \in R^{n \times m}$  can be reversed by multiplying each variable by its determined standard deviation and finally, by adding the mean, the residual matrix can be formed. The standard deviation and mean to be used here are determined from Eq. 6:

$$E = X - \hat{X} \quad (10)$$

The residual matrix  $E$  captures the variations in the observation space spanned by the loading vectors associated with the  $m - a$  smallest singular values (Chiang et al. 2001). This residual matrix will be used later in Eq. 12 to finally add the remaining variation of  $X$  not captured by the one retained score vector. If faults are then added to the one score vector capturing the correlation structure,  $T_{faulty} \in R^{n \times a}$  is formed, and then the faulty data can be acquired by

$$\hat{X}_{faulty} = T_{faulty}P^T \quad (11)$$

Then this  $\hat{X}_{faulty}$  is reverse normalized and finally the residual of Eq. 10 is added:

$$X_{faulty} = \hat{X}_{faulty} + E \quad (12)$$

Thus, the faulty data has been generated. The reason for setting  $a = 1$  is that the score vectors are orthogonal and ordered by the amount of variance:  $Var(t_1) \geq Var(t_2) \geq \dots \geq Var(t_a)$ . Thus, the work-around is to set  $a = 1$ , add faults to the one score vector, and then compute the  $\hat{X}_{faulty}$  and finally add the remaining variations captured in the residual matrix  $E$ . The faults added to the one score vector can be a fixed number subtracted or added for few intervals, a gradual evolving fault, random noise or maybe even other measures. The types of faults added will be discussed in the results section.

## Results

The data from the building follows certain patterns. That is, around 7.00 in the morning the temperature measurements in the four rooms starts to rise as does the CO2 ppm concentration. Accordingly, the radiator valves close (due to higher temperature) and the desired valve opening of the ventilation unit increases (due to higher CO2 concentration). The valve opening of the ventilation unit then accordingly also increases. Around 16.00 in the afternoon these sensors fall back to evening/nighttime operating conditions with no or almost no occupancy. Also, on weekends (Saturday, Sunday and other holidays), the

variables do not follow the same patterns as in the working days (Monday to Friday), but seem to follow patterns of no or almost no occupancy.

Three different artificial fault cases, produced by the PCA-based AFG method, will be introduced and run in ACCEPT; One fault corresponding to an open window during night, one during day and finally a ventilation fault during day. The reason for introducing both a daytime and a night time open window fault is that naturally more operations such as ventilation valve opening occur in the daytime. This ultimately translates to a more direct effect on a broader range of variables. It is thus anticipated that ACCEPT will be more capable of detecting an “open window fault” during the daytime than at nighttime since more variables will provide its indication. The ventilation fault is introduced since Fig. 2 states that “Dampers not working properly” is one of the typical faults. The open window fault is not directly related to this table, but will be considered more energy consuming than the duct leakage due to its inherent nature. The open window fault generation is instead justified by Fig. 1 where it is revealed that 16% of total energy consumption comes from space heating - thus faults in the heating system can be costly.

#### Open windows during night

The faulty data set is generated from the  $21 \times 25921$  dimensional training data set by running through the PCA-based AFG method.  $T_{faulty} \in R^{n \times a}$  is generated by subtracting a fault parameter of 30 from the values in the one score vector (of length 25921) 12 times of duration 108 time steps (9 h), on random weekdays. Since the fault is introduced during night, the beginning of the fault is at 22:00 and ending is 07:00. The fault parameter of 30 represents an arithmetic adjustment from the retained score vector reflecting the main correlation trend in the data set. The vector will thus increase in variance along a straight line (since PCA produces linear principal components). The unit of the score vector values in the score space is not easily interpretable but is revealed when projecting to the observation space. Here we have a “line” spanning itself in 21 dimensions and the unit of the slope of the line would involve 21 factors. Each point on the line would consist of 21 numbers readily interpretable in physically understandable units. A hypothetical unit of the slope could be °C pr. (ppm \* %) if we only had three dimensions. Setting the parameter to 30 is intended to serve as a proxy for an abrupt fault evolution, and can be characterized by an equivalent temperature- and a CO<sub>2</sub>-level drop and accompanying reactions by the different valve openings which preserve the correlations. This means that when CO<sub>2</sub> and temperature drops, the ventilation valves close and the radiator valves open (due to low temperature). In the ventilation fault case documented later in this paper a gradual introduction will be implemented by letting the fault parameter be a vector of length 108 (9 h duration).

Since the fault is induced on the temperature variables (and also on the CO<sub>2</sub>-level) in each of the four rooms, the target variable can be randomly chosen between those four rooms’ temperature variables. However, based on the contribution plot, it was found that variable number 10 contributes the most. Thus, this variable is selected as the target parameter to be used in ACCEPT. To generate a validation data set, another random 12 weekdays is chosen, thus making a data set that is not identical but drawn from nearly the same distribution as the test data set. The temperature drops below 20 °C and sometimes even below 17 °C.

To establish the ground truth for this variable, KDE is used on the training data, and on a 99% confidence interval the lower limit is approximately 19.95 °C. Thus, a value below 19.95 °C is set to correspond to an adverse event. The results generated by ACCEPT are shown in Fig. 4. As seen, the MDR is slightly lower than FAR, and in all cases the NMSE is higher compared to the NMSE of the benchmark case of another recent ACCEPT study (Egedorf and Shaker 2017). However, the results look good enough with PT and OT having MDR=0.9%, FAR=1.58% and DT=657 when using LIN and PT or OT so the fidelity is within an acceptable range. Detection performance is also acceptable since AUC is close to 1 in all cases. A figure will be shown later on the ventilation fault case showing the AUC of the ROC curve as an example of such a plot.

### Open window during day

Variable 10 is selected as the target parameter, since the contribution plot shows that this variable is contributing the most. Since the variable chosen here is the same as in the previous case, the same ground truth can be used. The ACCEPT results are shown in Fig. 5, and as seen ACCEPT is slightly better at detecting the introduced fault during day time compared to the night time fault case. Of course another thing that could explain the

Regression results...		
	lin	elm
Global Optimum	1.0000	36.0000
Optimized Values	0.6813	0.9817
Missed detection results...		
	lin	elm
Redline - Training	0.0097	0.0090
Redline - Validation	0.0090	0.0090
Predictive - Training	0.0090	0.0090
Predictive - Validation	0.0090	0.0090
Optimal - Training	0.0090	0.0090
Optimal - Validation	0.0090	0.0090
False alarm results...		
	lin	elm
Redline - Training	0.0160	0.0159
Redline - Validation	0.0259	0.0276
Predictive - Training	0.0158	0.0160
Predictive - Validation	0.0254	0.0272
Optimal - Training	0.0158	0.0159
Optimal - Validation	0.0596	0.0697
Detection time results...		
	lin	elm
Redline - Training	657.0000	657.0000
Redline - Validation	659.0000	659.0000
Predictive - Training	657.0000	657.0000
Predictive - Validation	659.0000	659.0000
Optimal - Training	657.0000	657.0000
Optimal - Validation	838.0000	838.0000

**Fig. 4** ACCEPT results on MDR, FAR and DT for fault case “Open-windows-during-night”. Regularization coefficient is 1.0 and number of hidden neurons is 36. Missed Detection Rate (MDR) is shown first, then False Alarm Rate (FAR) and finally Detection Time (DT)

Regression results...		
	lin	elm
Global Optimum	1.0000	36.0000
Optimized Values	0.6813	0.9817
Missed detection results...		
	lin	elm
Redline - Training	0.0097	0.0090
Redline - Validation	0.0090	0.0090
Predictive - Training	0.0090	0.0090
Predictive - Validation	0.0090	0.0090
Optimal - Training	0.0090	0.0090
Optimal - Validation	0.0082	0.0060
False alarm results...		
	lin	elm
Redline - Training	0.0156	0.0157
Redline - Validation	0.0253	0.0266
Predictive - Training	0.0156	0.0158
Predictive - Validation	0.0243	0.0266
Optimal - Training	0.0155	0.0156
Optimal - Validation	0.0517	0.0618
Detection time results...		
	lin	elm
Redline - Training	477.0000	477.0000
Redline - Validation	479.0000	479.0000
Predictive - Training	477.0000	477.0000
Predictive - Validation	479.0000	479.0000
Optimal - Training	477.0000	477.0000
Optimal - Validation	658.0000	658.0000

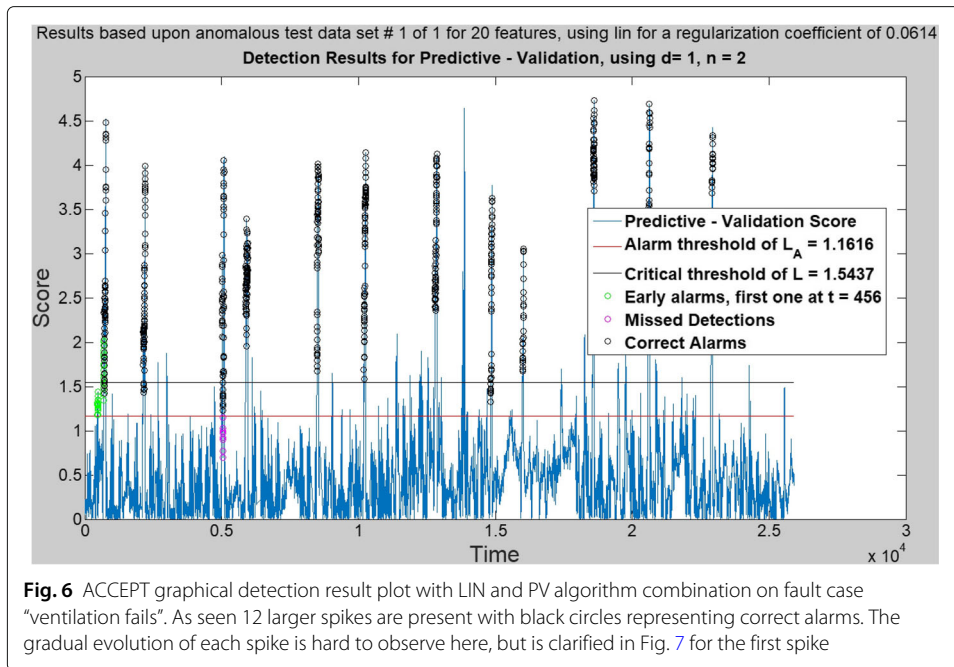
**Fig. 5** ACCEPT results on MDR, FAR and DT for fault case “Open-windows-during-day”. Regularization coefficient is 1.0 and number of hidden neurons is 36. Missed Detection Rate (MDR) is shown first, then False Alarm Rate (FAR) and finally Detection Time (DT)

slightly different results is that the testing and validation data sets are not the same in the two cases.

The daytime fault is like the night time case introduced on a random weekday every week for 12 weeks. However, here the fault is introduced during day time where people are occupying the rooms from 7:00 to 16:00. The validation data set is seeded as in the night time case (12 random weekdays other than those used in the test data set). The best algorithm combination seems to be LIN and PT with MDR=0.9%, FAR=1.56% and DT=477. In the next fault case there is a gradually evolving fault, which should be harder for ACCEPT to detect (higher MDR and FAR).

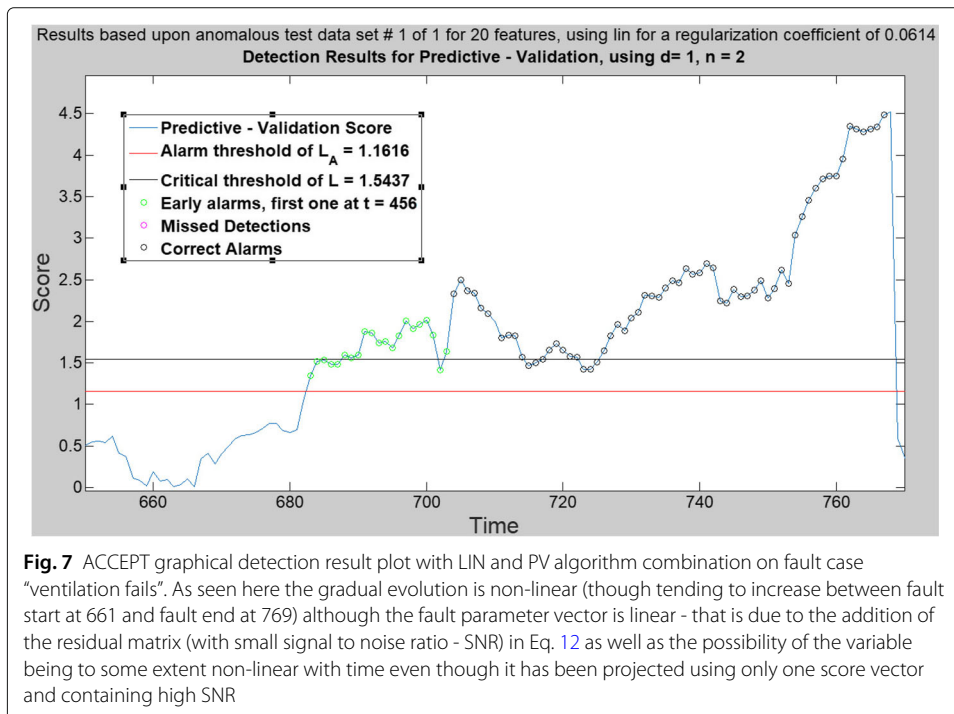
#### Ventilation fails during daytime

According to Fig. 2 a typical fault could be “dampers not working properly”. Thus, a fault case where ventilation valves fail is used by proxy. The simulation is run at daytime from 7:00 to 16:00 (9 h as in the previous cases) every week for 12 weeks (with random weekday selection). As mentioned previously, here the fault parameter of the PCA-based AFG method is a vector of 108 values (9 h) peaking at value 108 with a value of -30 (the slope



is negative to make the CO<sub>2</sub>-level spikes positive) - this creates gradually evolving spikes in the CO<sub>2</sub>-level in the four rooms, see Figs. 6 and 7.

Note here, that the fault parameter vector peak value is -30; the same as the value used in the open window cases. This makes the CO<sub>2</sub>-level peak at approximately 1600 ppm 12 times - peaking above or below 1600 ppm depending on which of the 12 faults is considered. Since the nature of a fault means that the correlation structure is not necessarily





preserved, the ventilation valves are set to be completely closed (since it is simulated that they fail) apart from what the PCA-based AFG actually dictates.

The variable mostly contributing, as determined by the contribution plot, is variable number 3 (CO<sub>2</sub>-level in ppm for room Ø22-508-1). Thus, the ground truth is established for this variable. According to (Prill 2013) the CO<sub>2</sub>-concentration should not exceed 1030 ppm inside buildings. Thus, in a sense, a ground truth defining an adverse event to correspond to a CO<sub>2</sub>-concentration above 1030 ppm could be used. However, the determination of the ground truth value can also, as previously used, be determined using confidence intervals. Here a value of approximately 742 ppm corresponds to the upper limit of a 99% confidence interval. However, this is known to be too low, and thus we chose a ground truth value of 1030 ppm derived from (Prill 2013). The ACCEPT results are shown in Fig. 8. As seen, the FAR is higher than MDR. However, it is acceptable since LIN and PV seem to be the best combination with MDR=1.15%, FAR=3.2% and DT=247. As expected as noted earlier this fault case is clearly harder for ACCEPT to detect (higher MDR and FAR). The regression fidelity is better than the two previous cases with NMSE values of 0.4669 and 0.8446 for LIN and ELM respectively (resulting

Regression results...		
	lin	elm
Global Optimum	0.0614	21.0000
Optimized Values	0.4669	0.8446
Missed detection results...		
	lin	elm
Redline - Training	0.0115	0.0159
Redline - Validation	0.0115	0.0144
Predictive - Training	0.0101	0.0159
Predictive - Validation	0.0115	0.0130
Optimal - Training	0.0101	0.0159
Optimal - Validation	0.0101	0.0115
False alarm results...		
	lin	elm
Redline - Training	0.0337	0.0426
Redline - Validation	0.0328	0.0441
Predictive - Training	0.0339	0.0420
Predictive - Validation	0.0320	0.0450
Optimal - Training	0.0338	0.0421
Optimal - Validation	0.0390	0.0492
Detection time results...		
	lin	elm
Redline - Training	247.0000	611.0000
Redline - Validation	247.0000	611.0000
Predictive - Training	247.0000	611.0000
Predictive - Validation	247.0000	611.0000
Optimal - Training	247.0000	611.0000
Optimal - Validation	247.0000	611.0000

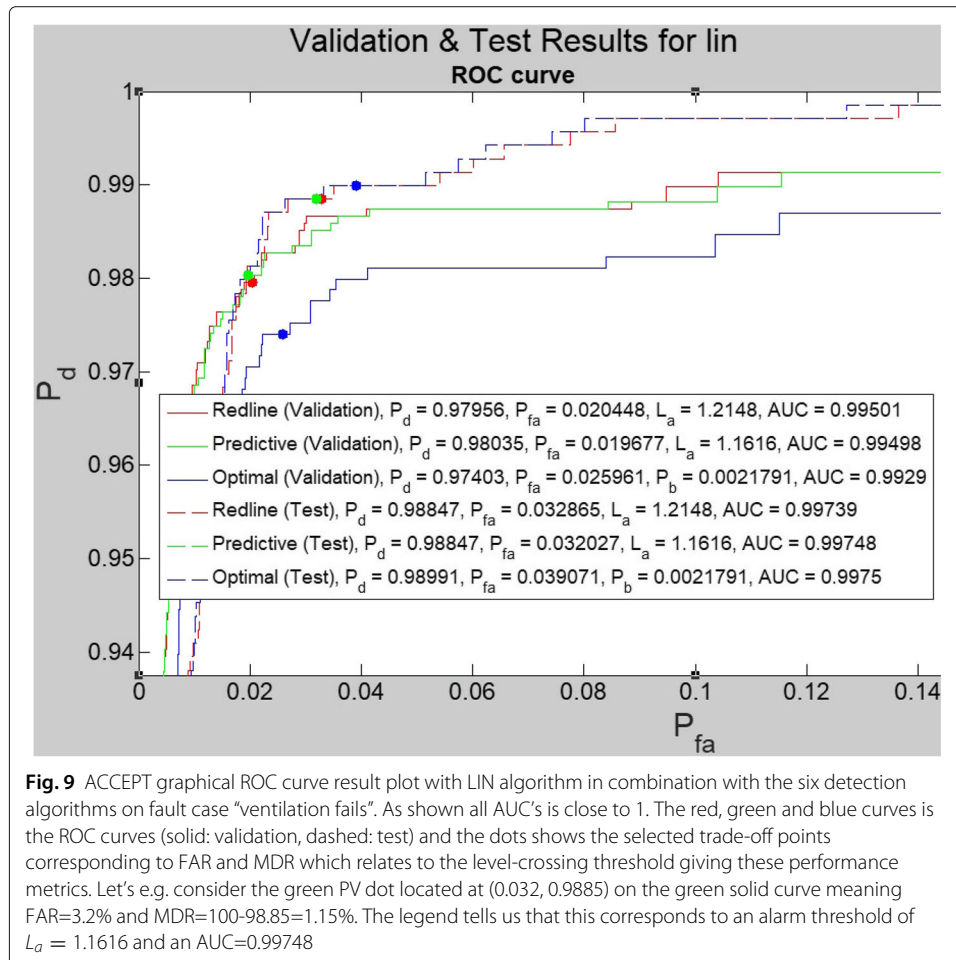
**Fig. 8** ACCEPT results on MDR, FAR and DT for fault case “ventilation fails”. Regularization coefficient is 0.0614 and number of hidden neurons is 21. Missed Detection Rate (MDR) is shown first, then False Alarm Rate (FAR) and finally Detection Time (DT)

from hyperparameters of 0.0614 and 21). Detection performance is also acceptable since AUC is close to 1 (see Fig. 9).

**Comparison of ACCEPT and CVA with KDE**

Since CVA is a dimensionality reduction technique (like PCA), it requires a parameter to select how many canonical variates can optimally be retained for the data set under consideration. This parameter is  $r \in N^+$  and different methods can be used to select the optimal. Two other parameters that need to be determined are the past and future lags,  $p$  and  $f$ . These lags are used to expand the observation matrix generating a past and a future matrix (see (Ruiz-Cárcel et al. 2015) for details) and the purpose is to take into account serial correlations between measurements of the same variable taken at different time instances. Lower/higher determined  $p$  and  $f$  values corresponds to the data being correlated with itself for shorter/longer time periods.

In this work the same approach to the one used in (Ruiz-Cárcel et al. 2015) is used to determine these mentioned parameters. The lags are determined by computing the autocorrelation function (ACF) of a stationary segment in the training data. Since the data are multivariate the sum of squares of each observation in the data are used to acquire a single signal for the ACF. To secure stationarity when computing the ACF the KPSS test



**Fig. 9** ACCEPT graphical ROC curve result plot with LIN algorithm in combination with the six detection algorithms on fault case “ventilation fails”. As shown all AUC’s is close to 1. The red, green and blue curves is the ROC curves (solid: validation, dashed: test) and the dots shows the selected trade-off points corresponding to FAR and MDR which relates to the level-crossing threshold giving these performance metrics. Let’s e.g. consider the green PV dot located at (0.032, 0.9885) on the green solid curve meaning FAR=3.2% and MDR=100-98.85=1.15%. The legend tells us that this corresponds to an alarm threshold of  $L_a = 1.1616$  and an AUC=0.99748

(Kwiatkowski et al. 1992) is used. Several stationary segments were found in the data and used in the analysis of the ACF and finally the lags are determined to be  $p = f = 2$ .

As mentioned, different methods have been suggested to select the value of  $r$ . The dominant singular values in a matrix  $D$  (see (Ruiz-Cárcel et al. 2015)) can be considered, but however as was found in (Ruiz-Cárcel et al. 2015) this can lead to an unrealistic model if the singular values decrease slowly. Therefore the method is to split the training data set and use one set as training set and the other for testing set. Then CVA is computed on different combinations of these split data sets as training and testing data sets using a range of values of  $r$ . The value of  $r$  is selected to be the one minimizing the false alarm rate. After several analysis testing different combinations of the data sets and using a range of  $r$  finally  $r = 16$  was found to be an optimal choice.

Using these parameters and the  $T^2$  metric as an indicator, the performance metrics are shown in Table 1. The Q metric represents the variation error in the residual space, where  $T^2$  represents the variation in the retained space. They are complementary, but in this case the Q health indicator performed poor compared to  $T^2$  - therefore  $T^2$  was selected for comparison. The reason for the use of KDE instead of fitting the  $T^2$  to e.g. a Gaussian distribution is that the data is non-linear.

In Table 1, OWDN, OWDD and Vent correspond respectively to Open Window During Night, Open Window During Day and Ventilation fault cases. As can be seen the performance metrics of MDR and FAR are similar in the first two cases while ACCEPT has a much lower MDR in the Vent fault case. What should be noted, as in another recent study on ACCEPT (Egedorf and Shaker 2017), the MDR of ACCEPT should in reality be higher in some cases when this comparison is done. ACCEPT is not using fault start and stop times (which CVA does) to compute the performance metrics but uses instead, among other things, the ground truth function. Thus, in the gradual evolving Vent fault case ACCEPT does not detect all data points in faulty region but still does not consider them as missed detection which CVA does, making the ACCEPT MDR much lower than that of CVA - see Fig. 7 where few data points after fault start at 661 to 683 are not considered missed detections. A quick estimate would then suggest that the ACCEPT MDR should be around  $(683-661)/108=20.37\%$  when compared to CVA. This suggest that the real performance metrics MDR and FAR of ACCEPT and CVA are quite similar. The detection time of CVA in Table 1 is taken as the time steps after fault start that a fault is detected. In the Vent fault case it is seen to be 27 time steps or 135 min. Thus CVA does not predict as ACCEPT does and therefore the DT of ACCEPT is negative in that table. Considering Fig. 7 again the prediction happens at  $t = 456$  and the fault starts at  $t = 661$  and thus a more correct prediction from ACCEPT in that case would be  $661-456=205$  time steps

**Table 1** Comparison table

CVA with KDE			
Fault case	MDR (%)	FAR (%)	DT (5 min)
OWDN	0.00	0.85	0
OWDD	0.00	0.90	0
Vent	25.00	0.85	27
ACCEPT			
OWDN	0.9	1.58	-657
OWDD	0.9	1.56	-477
Vent	1.15	3.2	-247

(17.08 h) and not the  $703-456=247$  (20.58 h) that ACCEPT computes. However, even though the comparison is difficult due to the inherent differences in definitions of how the performance metrics is computed this discussion suggest that ACCEPT is powerful in detecting and predicting faults when compared to the state-of-the-art method CVA with KDE. The performance metrics of MDR and FAR are similar but ACCEPT makes a prediction which is of course powerful when compared to CVA with KDE.

## Conclusion

Adverse condition and critical event prediction is an important subject in a variety of applications and it is very closely related to the area of fault detection. ACCEPT is a MATLAB-based framework developed to compare the performance of different machine learning and early warning algorithms. ACCEPT tests and compares these algorithms according to their ability to predict adverse events in arbitrary time-series data from systems or processes. In this paper, ACCEPT has been used for fault detection and prediction in an actual commercial building. Through using KDE, PCA-based contribution plots, the data from the building has been treated and used in ACCEPT for fault detection and prediction. A novel method for artificial fault generation is introduced. The proposed method uses PCA and finds applications in different areas, and is also used to generate fault data for analysis purposes in this work. The results obtained from ACCEPT have been evaluated, discussed and compared with CVA and KDE in the paper, and it was concluded that ACCEPT is more powerful - especially because of the prediction capability.

## Endnotes

<sup>1</sup> data from the US Department of Energy (Energy U.S.D.o. 2011)

<sup>2</sup> <https://www.sdu.dk/en/service/vejviser/odense>

## Acknowledgments

The authors would like to thank Cristobal Ruiz Carcel and Yi Cao from Cranfield University for helping with fitting the CVA method to the building data.

## Funding

This work has been supported by the Innovation Fund Denmark for the project COORDICY.

## Availability of data and materials

Please contact authors for data requests.

## Authors' contributions

SE carried out the numerical implementation, studied the results and drafted most of the manuscript. HRS participated in the design of the study, analysis of the results and validation, he also helped to draft the manuscript. RAM helped with using ACCEPT and also validation of the results. BNJ helped with the design of the study and coordination. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Center for Energy Informatics, Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Campusvej, 5230 Odense M, Denmark. <sup>2</sup>Intelligent Systems Division, NASA Ames Research Center, Moffett Field, Mountain view, 94035-0001 CA, USA.

Received: 30 April 2018 Accepted: 23 July 2018

Published online: 14 August 2018

## References

- Akinci B, Garrett JH, Akin (2011) Identification of functional requirements and possible approaches for self-configuring intelligent building systems. National Institute of Standards and Technology
- Bolón-Canedo V, Sánchez-Marroño N, Alonso-Betanzos A (2015) Recent advances and emerging challenges of feature selection in the context of big data. *Knowl-Based Syst* 86:33–45. <https://doi.org/10.1016/j.knosys.2015.05.014>
- Brutsaert H, Gao Y, Rim Y, Xie Y (2016) Using nasa accept (adverse condition and critical event prediction toolbox) for adverse thermal events in a smart building. Department of Statistical Science, Cornell University, Malott Hall, Ithaca
- Chiang LH, Russell EL, Braatz RD (2001) *Fault Detection and Diagnosis in Industrial Systems*. Springer, London
- Egedorf S (2017) Adverse condition and critical event prediction in smart commercial buildings. Master's thesis, University of Southern Denmark
- Egedorf S, Shaker HR (2017) Adverse condition and critical event prediction in cranfield multiphase of facility. In: 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC). IEEE, Shanghai. pp 557–564
- Energy U.S.D.o. (2011) Commercial Energy End-Use Splits, by Fuel Type (Quadrillion Btu). D and R International, Ltd, Silver Spring
- Government TD (2009) Strategi for Reduktion Af Energiforbruget i Bygninger. The Danish Government, Copenhagen
- Jørgensen BN, Kjærgaard MB, Lazarova-Molnar S, Shaker HR, Veje CT (2015) Challenge: Advancing energy informatics to enable assessable improvements of energy performance in buildings. In: Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems. ACM, Bangalore. pp 77–82
- Kothamasu R, Shi J, Huang SH, Leep HR (2004) Comparison of selected model evaluation criteria for maintenance applications. *Struct Health Monit* 3(3):213–224
- Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J Econ* 54(1-3):159–178
- Lazarova-Molnar S, Shaker HR, Mohamed N, et al (2016) Fault detection and diagnosis for smart buildings: State of the art, trends and challenges. In: Big Data and Smart City (ICBDSC), 2016 3rd MEC International Conference on Big Data and Smart City. IEEE, Muscat. pp 1–7
- Li DC, Lin WK, Lin LS, Chen CC, Huang WT (2017) The attribute-trend-similarity method to improve learning performance for small datasets. *Int J Prod Res* 55(7):1898–1913
- Lombard LP, Ortiz J, Pout C (2007) A review on buildings energy consumption information. *Energy Build* 40:394–398
- Martin RA, Santanu D, Janakiraman VM, Hosein S (2015) Accept: Introduction of the adverse condition and critical event prediction toolbox. Technical Report NASA/TM2015-218927, National Aeronautics and Space Administration (NASA)
- Mills E (2011) Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the united states. *Energy Efficiency* 4(2):145–173
- Odiowei P, Cao Y (2009) Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *Comput Aided Chem Eng* 27:1557–1562
- Odiowei P-EP, Cao Y (2010) Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *IEEE Trans Ind Inform* 6(1):36–45
- Okun O (2011) Feature Selection and Ensemble Methods for Bioinformatics: Algorithmic Classification and Implementations. IGI Global, Hershey PA
- Prill R (2013) Measuring carbon dioxide inside buildings - why is it important? Washington State University, Energy Program:1–4
- Rodríguez PVJ, Negrea M, Arkkio A (2008) A simplified scheme for induction motor condition monitoring. In: *Mechanical Systems and Signal Processing*. Vol. 22, Issue 5. Elsevier. pp 1216–1236
- Roth KW, Westphalen D, Feng MY, Llana P, Quartararo L (2005) Energy Impact of Commercial Building Controls and Performance Diagnostics: Market Characterization, Energy Impact of Building Faults and Energy Savings Potential. TIAX LLC, Lexington Massachusetts
- Ruiz-Cárcel C, Cao Y, Mba D, Lao L, Samuel R (2015) Statistical process monitoring of a multiphase flow facility. *Control Engineering Practice* 42:74–88
- Sha M, Wang A, Chu M, Liu L (2013) A virtual sample generation approach for blast furnace fault diagnosis. *Appl Mech Mater* 303–306:1379–1384
- Shaker HR, Lazarova-Molnar S (2017) A new data-driven controllability measure with application in intelligent buildings. *Energy and Buildings* 138:526–529
- Tommerup H, Kragh J, Svendsen S, Garn CB, Clausen AU, Riis P (2004) Energibesparelser i Eksisterende Og Nye Boliger. DTU Byg, Copenhagen
- Tuv E (2009) Feature Selection with Ensembles, Artificial Variables, and Redundancy Elimination. *Journal of Machine Learning Research* 10(February):1341–1366. <https://doi.org/10.1145/1577069.1755828>
- Zhang Y, Qiu J, Liu G, Zhao Z (2015) Fault sample generation for virtual testability demonstration test subject to minimal maintenance and scheduled replacement. *Math Probl Eng* 2015:1–8. <https://doi.org/10.1155/2015/645047>