

# NAPREDNE METODE ZAZNAVANJA NAPAK NA OSNOVI PODATKOV PRI DELOVANJU SISTEMA ZA KLIMATIZACIJO

Goran Andonovski, Aljaž Stare, Igor Škrjanc

## Izveček:

Pravočasno zaznavanje napak na industrijskih procesih je ključnega pomena za pravilno in optimalno delovanje celotnega sistema. Napake so lahko različnega tipa in različne narave ter se lahko pojavljajo na različnih sestavnih komponentah sistema. V tem članku je predstavljeno zaznavanje napak na klimatskih sistemih s pomočjo dveh naprednih metod na osnovi toka podatkov, in sicer z analizo glavnih komponent (PCA) in s samorazvijajočo se metodo na osnovi mehkega modela. Samorazvijajoča se metoda v sprotnem načinu zgradi različne podmodele oziroma detektorje napak za posamezno napako. Nato lahko s pomočjo teh podmodelov zaznavamo in identificiramo (izoliramo) napako. Metodi smo testirali na realnih podatkih, pridobljenih iz realnega klimatskega sistema. Prednost naprednih metod je ta, da lahko napako zaznamo, še preden ta vpliva na delovanje celotnega sistema.

## Ključne besede:

zaznavanje napak, PCA, samorazvijajoči se mehki model, klimatski sistem

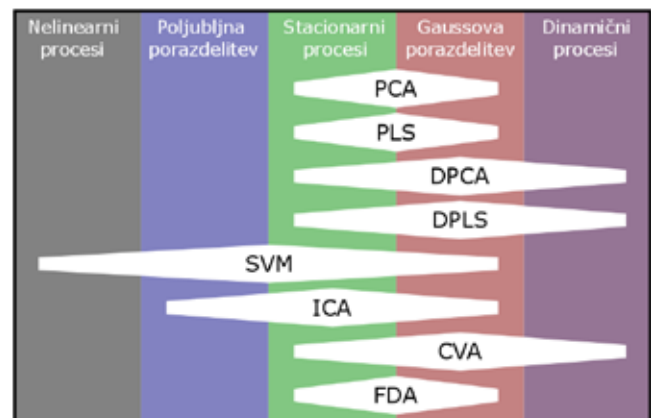
## 1 Uvod

Sodobni industrijski obrati so zelo kompleksni in sestavljeni iz različnih dinamičnih podprocesov. Pogosto obratujejo v različnih obratovalnih pogojih in pod spremenljivimi zunanji vplivi. Ne glede na to, ali gre za zelo velike in kompleksne sisteme ali pa za majhne in enostavne, so zahteve v industriji za varnejše in zanesljivejše delovanje vedno večje. Pri tem imajo pomembno vlogo metode spremljanja procesov oziroma metode zaznavanja napak. Klasične metode spremljanja procesov temeljijo le na spremljanju posameznih (merljivih) spremenljivk, če so znotraj predhodno opredeljenih meja (metoda limitnih vrednosti). Vendar spremljanje procesov lahko znatno izboljšamo, če upoštevamo več merjenih spremenljivk, ki so nam na voljo, in sicer v kombinaciji z naprednimi in inteligentnimi metodami.

Na področje zaznavanja napak spadajo metode, ki lahko odkrijejo, če se je pojavila napaka v sistemu, in določajo čas, v katerem se je napaka pojavila. Metode zaznavanja napak lahko razdelimo v tri večje skupine:

1. metode na podlagi matematičnih modelov procesa [1, 2],
2. statistične metode [3-5] ter
3. metode na podlagi podatkov [6-8].

V prvo skupino spadajo metode, ki temeljijo na matematičnem modelu in zahtevajo osnovno znanje o fizikalnem ozadju procesa. Slabost teh metod je ta, da ne upoštevajo vpliva motenj, lezenja parametrov in uporabljajo le a priori znanje o napakah. Druga skupina metod temelji na statističnih analizah podatkov procesa [3]. Te metode uporabljajo shranjene podatke za pridobitev znanja o



Slika 1 : Prikaz nekaterih uveljavljenih statističnih metod in področje njihovega delovanja

**Dr. Goran Andonovski**, univ. dipl. inž., Univerza v Ljubljani, Fakulteta za elektrotehniko; **dr. Aljaž Stare**, univ. dipl. inž., Metronik, d. o. o., Ljubljana; **prof. dr. Igor Škrjanc**, univ. dipl. inž., Univerza v Ljubljani, Fakulteta za elektrotehniko

procesu, ko je ta v normalnem delovanju oziroma ko se pojavi napaka. Najosnovnejše metode so: analiza glavnih komponent [9, 10], metoda delnih najmanjših kvadratov [11], metoda neodvisnih komponent [5] itn. Nekaj teh statističnih metod je prikazanih na *sliki 1*. Opazimo, da so te metode v glavnem namenjene za stacionarne procese z normalno porazdelitvijo.

V tretjo skupino spadajo metode, ki uporabljajo tok podatkov v kombinaciji s statističnimi orodji za sprotno analizo podatkov. Mi se bomo še natančneje posvetili metodam s samorazvijajočo se strukturo in adaptivnimi parametri. Te metode so primerne za nelinearne in za dinamične sisteme. Metoda za zaznavanje napak na podlagi samorazvijajočega se deljenja prostora na Gaussove roje je predstavljena v [6]. V [7] je predstavljena metoda za zaznavanje napak, ki temelji na načelu TEDA (ang. *typicality and eccentricity data analytics*) [8].

V tem prispevku bomo najprej predstavili obravnavani klimatski sistem in potencialne napake, ki se lahko pojavijo v sistemu tako na izvršnih členih kot tudi na senzorjih. V eksperimentalnem delu prispevka smo primerjali dve metodi zaznavanja napak. Prva metoda, ki spada v skupino statističnih metod, je metoda glavnih komponent (PCA). Metoda je uporabna, ko imamo opravka z večdimenzionalnimi podatki, saj omogoča predstavitev podatkov z manjšim številom spremenljivk. Druga metoda uporablja statistične in dinamične informacije o procesu, ki jih pridobi s sprotim učenjem na osnovi toka podatkov (ang. *data stream*) v kombinaciji z mehanizmom samorazvijanja. To nam omogoča bolj fleksibilno prilagajanje parametrov metode in na ta način lahko zajamemo nestacionarne in dinamične procese.

## 2 Klimatski sistemi - HVAC

Klimatski sistemi za pripravo zraka (ang. *Heating, Ventilation and Air Conditioning*, HVAC) so del skoraj vsakega industrijskega objekta. Tipični ele-

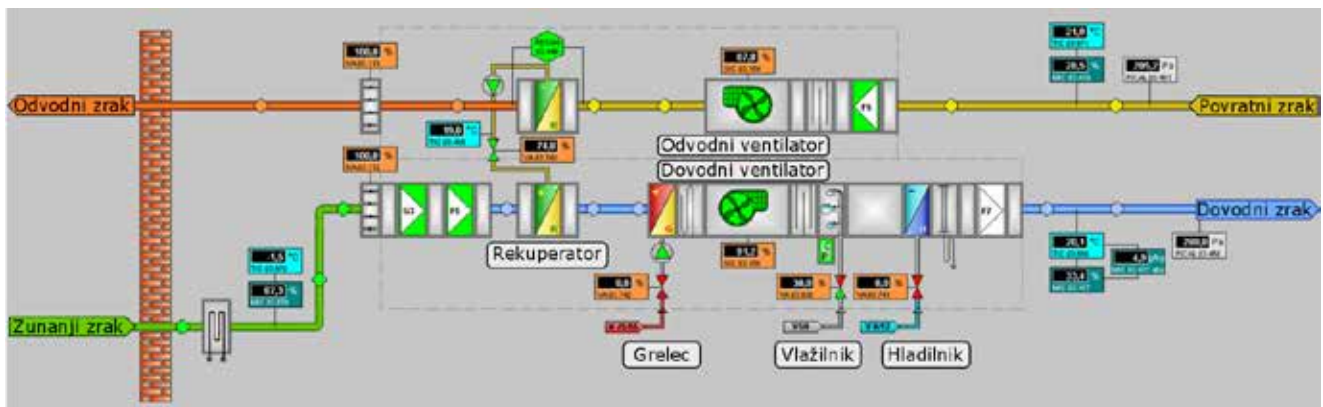
menti sistema za pripravo zraka so ventilatorji, cevi, ventili, grelni in hladilni elementi ter mešalniki zraka. Shema tipičnega klimatskega sistema HVAC je predstavljena na *sliki 2*. Za podani sistem smo pridobili merjene podatke v obdobju sedmih mesecev (junij 2016–december 2016). Pridobljene podatke smo uporabili za nastavljanje in optimizacijo detektorjev napak.

### 2.1 Možne napake na sistemu HVAC

V tem podpoglavju so predstavljene tipične napake, ki se lahko pojavijo na klimatskem sistemu. Za potrebe testiranja v poglavju 3 so generirani testni podatki (simulacijski teki), kjer se v različnih časovnih intervalih pojavljajo različne napake v delovanju sistema, ki naj jih detektorji uspešno zaznavajo in prepoznajajo. Predvidene in možne so napake na različnih signalih, kot jih prikazuje *tabela 1*.

**Tabela 1:** Seznam signalov na sistemu HVAC

	Opis signala
AKTUATORJI	Ventil rekuperatorja
	Ventil grelca
	Ventil vlažilca
	Ventil hladilnika
	Dovodni ventilator
	Odvodni ventilator
	Lopute dovodnega zraka Lopute odvodnega zraka
SENZORJI	Temperatura dovodnega zraka
	Temperatura odvodnega zraka
	Temperatura povratnega zraka
	Temperatura rekuperacije
	Temperatura grelnega medija
	Temperatura hladilnega medija
	Vlažnost dovodnega zraka
	Vlažnost odvodnega zraka
	Vlažnost povratnega zraka
	Tlak/pretok dovodnega zraka
	Tlak/pretok odvodnega zraka
	Temperatura dovodnega zraka



**Slika 2:** Shema realnega sistema HVAC

Predvidimo tudi več vrst napak na zgoraj navedenih signalih, in sicer:

- ▶ zakasnitev signala,
- ▶ premik signala po amplitudi,
- ▶ izpad komunikacije ter
- ▶ odpoved elementa.

### 3 Zaznavanje napak na sistemu HVAC

V tem poglavju bo opisan postopek zaznavanja napak z dvema metodama: PCA in metodo na osnovi oblakov z mehanizmom samorazvijanja. Obe metodi potrebujeata dve množici podatkov za vsako napako. Prednost druge metode je ta, da se nauči v sprotnem načinu (ang. *one-pass algorithm*) ter lahko zazna novo stanje, ki ga nismo predvideli. Za izgradnjo in testiranje detektorjev napak potrebuje mo množico učnih in testnih podatkov.

#### 3.1 Metoda glavnih komponent - PCA

Za gradnjo detektorjev različnih napak smo preizkusili analizo glavnih komponent. Metoda PCA (ang. *Principal component analysis*) poišče linearna razmerja med spremenljivkami ter omogoči preslikavo osnovnih podatkov v prostor glavnih komponent, katerega osi so med seboj ortogonalne. Drugače povedano: metoda PCA definira transformacijsko matriko  $\mathbf{P}$ , ki preslika vhodne podatke  $\mathbf{X}$  v matriko zadetkov  $\mathbf{T}$  [12]. Pravzaprav gre za transformacijo (rotacijo) osi koordinatnega sistema v nov ortogonalni sistem, ki rezultira v največjo kovarianco podatkov glede na novo izbrane osi. To lahko opišemo z enačbo:

$$\mathbf{T} = \mathbf{X}\mathbf{P} \quad (1)$$

Matrika glavnih vektorjev  $\mathbf{P}$  predstavlja statistični model procesa in nima realnega ozadja, saj je sestavljen iz psevdospremenljivk. Lahko sestavimo delno matriko (model) glavnih lastnih vektorjev  $\mathbf{P}_s$ , tako da vzamemo le stolpce matrike  $\mathbf{P}$ , ki nosijo večino variance oziroma informacije vhodne matrike  $\mathbf{X}$ . Na ta način dobimo aproksimirano matriko podatkov:

$$\mathbf{T}_s = \mathbf{X}\mathbf{P}_s \quad (2)$$

$$\mathbf{X}_s = \mathbf{T}_s\mathbf{P}_s^T \quad (3)$$

Z zanemaritvijo nekaj komponent smo izgubili del informacije o osnovnih meritvah. To lahko zapišemo z matriko ostankov (residualov)  $\mathbf{E}_x$ , ki jo imenujemo tudi prostor šuma. Sedaj lahko matriko podatkov zapišemo z naslednjo enačbo:

$$\mathbf{X} = \mathbf{X}_s + \mathbf{E}_x = \mathbf{T}_s\mathbf{P}_s^T + \mathbf{E}_x \quad (4)$$

Matriko šuma (residualov) bomo kasneje uporabili za zaznavanje napak. V [12] je metoda PCA predstavljena bolj podrobno.

#### 3.2 Samorazvijajoča se metoda

Metoda za zaznavanje napak, ki bo razložena v nadaljevanju, je zasnovana na osnovi samorazvijajočega se mehkega modela AnYa, ki temelji na izračunu lokalnih gostot (ang. *local density*). Model AnYa je sestavljen iz mehkih pravil:

$$\mathfrak{R}^i: \text{IF } (\mathbf{x}_k \sim X^i) \text{ THEN } \mathbf{x}_k \in \text{Class}^i \quad (5)$$

kjer je  $\mathbf{x}_k = [\mathbf{x}_k(1), \mathbf{x}_k(2), \dots, \mathbf{x}_k(m)]$   $m$ -dimenzionalni vhodni vektor. Operator  $\sim$  je izražen kot mera pripadnosti trenutnega vzorca  $\mathbf{x}_k$  k obstoječim oblakom  $X^i$ . Mera pripadnosti predstavlja lokalna gostota vzorca in se izračuna z naslednjo enačbo:

$$\gamma_k^i = \frac{1+T^i}{1+(\mathbf{x}_k - \mu_{M^i}^i)^T(\mathbf{x}_k - \mu_{M^i}^i) + T^i}, i = 1, \dots, c \quad (6)$$

kjer je  $\mu_{M^i}^i$  srednja vrednost  $i$ -tega oblaka.  $T^i$  je skalar in se izračuna z naslednjo enačbo:

$$T^i = \frac{M^i - 1}{M^i} \text{trace}(\mathbf{A}^i \Sigma_{M^i}^i). \quad (7)$$

V enačbi (7) je potrebno izračunati še kovariančno matriko podatkov  $\Sigma$ . V [13] so avtorji predlagali rekurziven način izračuna kovariančne matrike z naslednjim algoritmom:

$$M^i \leftarrow M^i + 1 \quad (8)$$

$$\mu_{M^i}^i \leftarrow \frac{M^i - 1}{M^i} \mu_{M^i}^i + \frac{1}{M^i} \mathbf{x}_k \quad (9)$$

$$\mathbf{S}_{M^i}^i \leftarrow \mathbf{S}_{M^i - 1}^i + (\mathbf{x}_k - \mu_{M^i - 1}^i)(\mathbf{x}_k - \mu_{M^i}^i)^T \quad (10)$$

$$\Sigma_{M^i}^i \leftarrow \frac{M^i - 1}{M^i} \Sigma_{M^i}^i \quad (11)$$

kjer so začetna stanja  $M^i = 0$ ,  $\mu_0^i = \mathbf{x}_1$  in  $\mathbf{S}_0^i = \mathbf{0}$ . V [14] je predstavljen podroben opis metode in način, kako jo uporabimo v praksi.

#### 3.3 Zaznavanje napak

Za gradnjo detektorja napak pri obeh metodah smo izbrali pristop, pri katerem se v koraku učnega detektorju podasta dve množici podatkov. Prva množica predstavlja delovanje sistema HVAC brez napak, druga množica podatkov pa predstavlja obdobje, ko je prisotna napaka, ki jo želimo zaznati. V fazi vrednotenja imamo eno množico podatkov, ki vsebuje področje normalnega delovanja kot tudi področje z napako.

##### 3.3.1 Faza učenja

V fazi učenja zgradimo detektor napake z uporabo dveh naborov podatkov. Z metodo PCA določimo transformacijsko matriko  $\mathbf{P}$  za oba nabora

podatkov. Potem sledi izbor glavnih komponent, v katerem izbrane komponente predstavljajo 95 % skupne variance (podrobno je opisano v [12]). V primeru samorazvijajočega se modela pa na osnovi obeh naborov podatkov zaznamo oblake, ki predstavljajo normalno delovanje procesa, in oblake, ki predstavljajo napako.

### 3.3.2 Faza vrednotenja

V fazi ovrednotenja, kot smo že omenili zgoraj, potrebujemo eno množico podatkov, ki vsebuje področje normalnega delovanja in prisotnost napake.

V primeru samorazvijajočega se modela na osnovi oblakov za vsak podatek  $x_k$  iz testnega seta podatkov izračunamo maksimalno gostoto oblakov brez napake in z napako. Z naslednjo funkcijo, ki je pravzaprav del posledičnega dela enačbe (2.4), lahko vsak podatek razvrstimo kot normalno delovanje ali napako:

$$Napaka = \begin{cases} 1, & \max_i \gamma_k^i(X_{F=0}^i) < \max_i \gamma_k^i(X_{F=1}^i) \\ 0, & \text{sicer} \end{cases} \quad (12)$$

kjer 1 pomeni, da smo odkrili napako, medtem ko 0 pomeni normalno obratovanje sistema.

Ovrednotenje obeh metod je bilo izvedeno s pomočjo matrike zamenjav kot ocena robustnosti metode:

$$ACC = \frac{TP+TN}{total\ samples} \times 100 [\%] \quad (13)$$

kjer je  $TP$  število pravilno zaznanih vzorcev in  $TN$  število napačno zaznanih vzorcev.

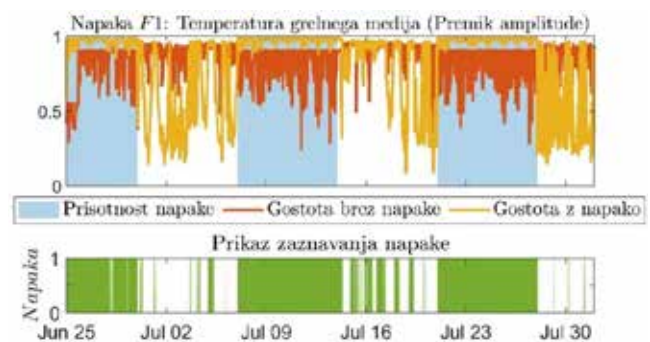
## 4 Rezultati

V eksperimentalnem delu smo izbrali tri različne napake, ki se lahko pojavijo na sistemu HVAC, kot je prikazano v tabeli 2. Za vsako napako smo s pomočjo simulatorja generirali učne in testne množice podatkov.

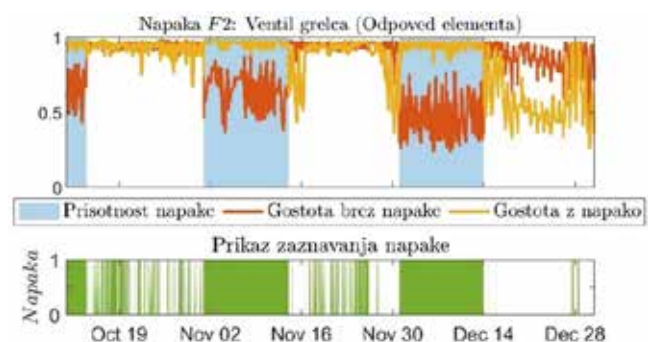
Na slikah 3, 4 in 5 so prikazani rezultati zaznavanja napak (testna množica podatkov) s pomočjo samorazvijajočega se modela na osnovi oblakov. Na teh slikah je najprej predstavljena primerjava med maksimalno gostoto oblakov brez napake in ma-

Tabela 2 : Seznam testiranih napak

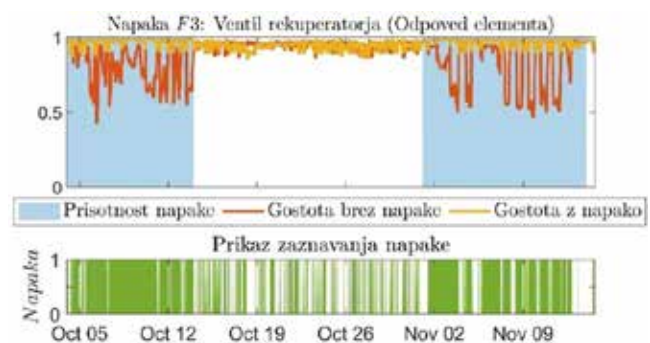
	Opis napake	Tip napake
F1	Temperatura grelnega medija	Premik amplitude
F2	Ventil grelca	Odpoved elementa
F3	Ventil rekuperatorja	Odpoved elementa



Slika 3 : Rezultati zaznavanja napake F1



Slika 4 : Rezultati zaznavanja napake F2



Slika 5 : Rezultati zaznavanja napake F3

ksimalno gostoto oblakov z napako. Nato z enačbo (12) določimo, ali je trenutni vzorec napaka ali normalno obratovanje. To prikazuje drugi graf na slikah.

V tabeli 3 so prikazani rezultati uspešnosti zaznavanja napak za obe metodi, kjer je eCB okrajšava za metodo na osnovi oblakov. Iz tabele 3 je razvidno, da samorazvijajoči se model dosega boljše rezultate in je primeren za uporabo na realnem sistemu.

Tabela 3 : Prikaz uspešnosti zaznavanja napak (ACC [%])

	F1	F2	F3
PCA	86.59	75.35	67.90
eCB	<b>89.03</b>	<b>75.76</b>	<b>69.04</b>

## 5 Zaključek

V tem delu smo predstavili možne napake, ki se lahko pojavijo na klimatskih sistemih, in dve različni metodi za zaznavanje teh napak. Prva metoda, metoda glavnih komponent, temelji na statistični analizi podatkov in transformaciji (preslikavi) prostora podatkov v prostor glavnih komponent. Druga metoda s sprotnim učenjem na osnovi podatkov pridobiva novo znanje o procesu in razdeli prostor na oblake podatkov. Obe metodi smo testirali na treh različnih napakah, ki se lahko pojavijo na sistemu HVAC. Na osnovi dobljenih rezultatov lahko sklepamo, da obe metodi dajeta približno podobne rezultate. Prednost druge metode je ta, da lahko v sprotnem načinu odkrije novo področje delovanje procesa in na ta način lahko dodamo novo znanje o samem procesu.

### Viri

- [1] Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical knowledge-based redundancy - A survey and some new results. *Automatica*, 26(3), 459-474.
- [2] Klančar, G., Juričić, Đ., & Karba, R. (2002). Robust fault detection based on compensation of the modelling error. *International Journal of Systems Science*, 33(2), 97-105.
- [3] Qin, S. J. (2003). Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17(8-9), 480-502.
- [4] Klančar, G., & Škrjanc, I. (2002). Metoda glavnih komponent pri odkrivanju in izolaciji napak: primer hidravličnega procesa in procesa fermentacije. *Elektrotehniški Vestnik*, 69(5), 311-316.
- [5] Zhang, Y., & Qin, S. J. (2008). Improved Non-linear Fault Detection Technique and Statistical Analysis. *AIChE Journal, Process Systems Engineering*, 54(12), 3207-3220.
- [6] Lemos, A., Caminhas, W., & Gomide, F. (2013). Adaptive fault detection and diagnosis using an evolving fuzzy classifier. *Information Sciences*, 220, 64-85.
- [7] Bezerra, C. G., Costa, B. S. J., Guedes, L. A., & Angelov, P. P. (2016). An evolving approach to unsupervised and Real-Time fault detection in industrial processes. *Expert Systems with Applications*, 63, 134-144.
- [8] Angelov, P. (2014). Anomaly Detection Based on eccentricity analysis. *IEEE Symposium on Evolving and Autonomous Learning Systems (EALS)*, 1-8.
- [9] Li, W., Yue, H., Valle-Cervantes, S., & Qin, S. J. (2000). Recursive PCA for adaptive process monitoring. *Journal of Process Control*, 10, 471-486.
- [10] Klančar, G. (2000). Fault Detection and Isolation by means of Principal Component Analysis. *Cybernetics & Informatics Eurodays: Young Generation Viewpoint : PhD Workshop*, 1-6.
- [11] Li, G., Qin, S. J., & Zhou, D. (2010). Geometric properties of partial least squares for process monitoring. *Automatica*, 46(1), 204-210.
- [12] Stržinar, Ž. (2017). Modeliranje in zaznavanje napak v klimatskih sistemih. Doktorska disertacija. University of Ljubljana.
- [13] Blažič, S., Angelov, P., & Škrjanc, I. (2015). Comparison of Approaches for Identification of All-data Cloud-based Evolving Systems. *2nd IFAC Conference on Embedded Systems, Computer Intelligence and Telematics CESCIT 2015*, 129-134.
- [14] Andonovski, G., Blažič, S., & Škrjanc, I. (2019). Evolving fuzzy model for fault detection and fault identification of dynamic processes. *Predictive Maintenance in Dynamic Systems: Advanced Methods, Decision Support Tools and Real-World Applications* (pp. 269-285).

## Advanced data-based methods for fault detection in HVAC systems

### Abstract:

Fault detection in industrial systems plays a crucial part of the correct and optimal operation of the different processes. The possible faults could be of different types and natures and occur on different components of the process. This paper presents fault detection on air conditioning systems using an advanced data-based method. The method evolves its structure in an online manner and builds various fault detectors for each fault. Furthermore, we could detect and identify each fault separately. The proposed method was tested on real data from a real air conditioning system. The advantage of the evolving method is the ability to detect new faults before it affects the operation of the process.

### Keywords:

advanced methods, fault detection, HVAC system