DOI: 10.18462/iir.gl2022.46

Correlation Analysis of evaporation pressure readings in CO2 supermarket refrigeration systems

Andreas SCHULTE *[(a)], Wilhelm TEGETHOFF[(a)], Juergen KOEHLER[(a)]

[(a)] TU Braunschweig, Institut für Thermodynamik, Hans-Sommer Straße 5
Braunschweig, 38106, Germany, ift@tu-braunschweig.de
*Corresponding author: andreas.schulte@tu-braunschweig.de

ABSTRACT

Digital Twins for refrigeration systems can be used to create advanced digital services. These services rely on correct and consistent data. A typical supermarket refrigeration system measures the evaporation pressure at multiple points. Because all the pressure sensors, on a given pressure level, are connected to a common suction line, they provide redundant measurements, that can be leveraged for fault detection. Furthermore, the pressure readings carry structural information about the layout of the system. In this work a method for fault detection using evaporation pressure readings in a CO2 supermarket refrigeration system is presented. Based on a correlation analysis of the evaporation pressures, the presented method can identify inconsistent measurements. Beyond that, this method can identify the components belonging to the individual pressure levels. The method is applied to real world data from supermarket monitoring systems. Examples of found inconsistencies are shown and the ability to identify structure is demonstrated.

Keywords: Supermarket, Fault detection, Structure Identification, Carbon Dioxide

1. INTRODUCTION

Industries around the globe are in the midst of a new revolution, the fourth industrial revolution. This brings forward the digitalisation of machines and factories. With the digitalization of refrigeration systems new opportunities for advanced digital services emerge. Many of these digital services riley on Digital Twins, for example fault detection or control optimization. With more than 100.000 supermarkets in Europe (European Commission, 2014) rolling out Digital Twins at that scale require automated methods to create the Digital Twins. For this it is essential to have methods for identifying the layout of a supermarket, as well as a way the provide early fault detection on the used data.

Figure 1 shows the basic layout of a standard transcritical CO2 Booster System. Starting at the compressors the refrigerant gets conveyed to the gas cooler (GC) on the outside of the store. Here the refrigerant rejects heat to the ambient. After the gas cooler the refrigerant get throttled by the high-pressure valve (HPV) into the receiver, where liquid and gaseous phase get separated. The gas phase gets throttled into the suction manifold of the compressor, via the flash gas bypass valve (FGB). The liquid refrigerant is supplied to the display cases is the store room. Here an expansion valve throttles the refrigerant down to the evaporation pressure. Depending on the display cabinet this is either the medium-temperature (MT) evaporation pressure or the low-temperature (LT) evaporation pressure. Respectively the display cabinets are connected to the MT or LT compressor.

The display cabinets are often controlled by an electronic controller, for example Danfoss AK-CC 550A or AK-CC 750A. These controllers in combination with sensors and an electronic expansion valve form a complete regulating unit for a display case, capable of controlling for example superheat, temperature and defrost (Danfoss, 2018). They can control one or more individual evaporators. For the calculation of the superheat each controller is equipped with a pressure sensor. In the case of multiple evaporators being controlled by one controller, one pressure sensor is used for all controlled evaporators (Danfoss, 2018). This means, that in a typical supermarket refrigeration system the evaporation pressure is measured at multiple locations, each individual controller has its own pressure sensor in addition to a pressure sensor is on the compressor rack.

Because all the evaporation pressure sensors, on a given pressure level, are connected by a common suction line, they provide redundant measurements, that can be leveraged for fault detection. Furthermore, the physical connection between the sensors provides a structural information about the layout of the supermarket system, that can be identified.

Most modern supermarkets use a monitoring system to supervise and log the operation of the supermarket. Here the measurements of the sensors are logged in regular intervals to a database. This database allows it to access historical data for the supermarket, upon which analysis can be performed. The interval for sensor data typically is in the order of fifteen minutes, but can also be as low as one minute or as high as several hours. The evaporation pressure readings in the monitoring system are often logged, not as the actual pressure value, but rather as the corresponding equilibrium temperature. Therefore, in the following, evaporating pressure and evaporating temperature are used interchangeably, and in the diagrams the evaporating temperature is shown.

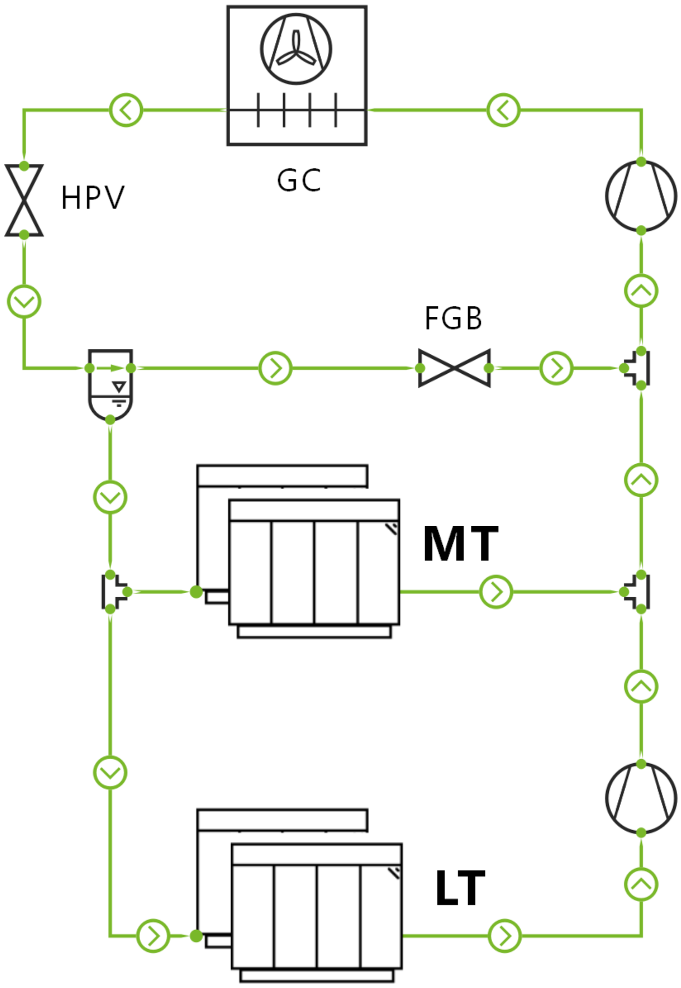

Figure 1: Principle Flow Chart for a transcritical CO2 Booster System used in supermarket refrigeration. Two evaporation pressure levels exist, the medium-temperature (MT) level and the low-temperature (LT) level.

The interaction between the compressors and the display cabinets in a supermarket refrigeration system result in a constantly fluctuating evaporation pressure in the system. Because of the physical connection between all pressure sensors on a distinct pressure level it is assumed that all pressure readings on this pressure level move in the same direction at the same time, hence are strongly correlated. Furthermore, it is assumed that the pressure readings on any other distinct pressure levels are uncorrelated. This assumption seems plausible in a supermarket system, since the fluctuation are not the result of a common process, but rather the result of random influences of the environment on the display cabinets. Therefore, the expectation is to see a strong correlation for the pressure readings of a distinct pressure level. In addition to that all pressure sensor on a pressure level should form a group of strongly correlated pressure readings. If the pressure readings are grouped by there correlations, the resulting clusters should represent the layout of the supermarket, meaning that the number of clusters should be equal to the number of pressure levels and the members of the clusters should be the evaporators belonging the respective pressure levels. A deviation from the expected number of pressure levels indicates a fault either in the refrigeration system itself or in the data logged in the monitoring system.

2. METHOD

The workflow presented in this work consists of the following steps:

1. Selection of pressure readings out of the data,
2. Calculation of the correlation matrix,
3. Clustering of all strongly correlated pressure readings,
4. Structure identification and fault detection.

2.1. Selection of pressure readings

As alluded to in the Introduction, the physical connection of the pressure sensors by the common suction line results in strongly correlated pressure readings. For analysing the pressure reading of a supermarket the relevant data first needs to be selected from the monitoring system. To make the data in the database of the monitoring useful for a human operator meta information in needed, to identify the variables. In the simplest case this can be a structured naming scheme. Meaning that in the simplest case selecting the evaporation pressures can be achieved via a name-based filter. As for the time duration used for analysis, care has to be taken, as short durations allow fast detection of changes in the data, but the method becomes more susceptible to noise. Long duration on the other hand increase the time until a change in the date is detected. With a typical sampling time between one and fifteen minutes, a duration of one week is proposed as good compromise between speed and robustness.

2.2. Calculation of the correlation matrix

The correlation coefficient (Zeidler, E., 2013) of two variables x and y is a measure of the linear correlation between these two variables. It is the covariance of the two variables, normalized by the product of their standard deviations. This normalization means that the resulting correlation coefficient can only be between 1 and -1. For a give sample of n variable of x and y with their mean values $\bar{x}$ and $\bar{y}$ the correlation coefficient is given by:

$$r_{x,y} = \frac{\frac{1}{n}\sum_{i=0}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=0}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=0}^{n}(y_i-\bar{y})^2}} \quad \text{Eq. (1)}$$

One limitation of the correlation coefficient is, that it can only reflect the linear correlation and ignores non-linear relationships between the variables. As non-linear correlations between different pressure readings are not expected,

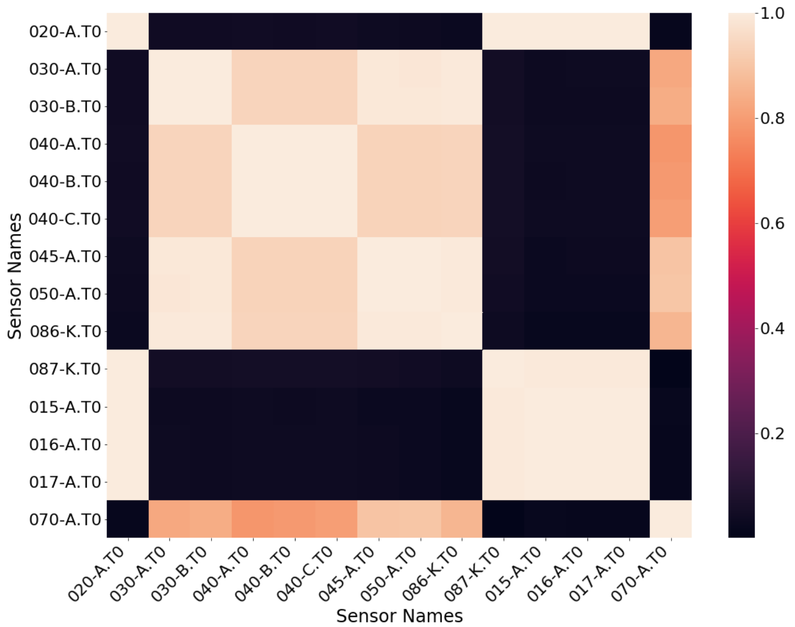

Figure 2: Correlation matrix of an exemplary supermarket. It can be seen, that the values all lay either near one or near zero.

this property of the correlation coefficient is seen as positive. When there are more than two variables the correlation coefficient is visualized in form of a matrix, the correlation matrix, where every entry describes the individual correlation between two variables.

Calculating the correlation matrix of a given dataset is done by calculating the pair-wise correlation coefficient Eq. (1) for every corresponding entry in the matrix. This is a typical function of many data analysis frameworks. In Figure 2 the correlation matrix for an exemplary supermarket is shown. The values of the individual correlation coefficient are translated into a colour scale for easier visual interpretation. It can be seen, that the correlation coefficients are either near zero or near one. This means, that the pressure reading is either near ideally correlated or totally uncorrelated.

2.3. Hierarchical clustering

In order to identify cluster of strongly correlated pressure readings out of the correlation matrix a clustering algorithm is used. The method for identifying the cluster of strongly correlated pressure readings used in this work is the hierarchical clustering. The basic idea of hierarchical clustering is to build a hierarchy of clusters based on the distance between their values. The distance measure is chosen in this work so that the resulting distance is small for correlated values and large for uncorrelated values. The result of this is a dendrogram (Figure 3) where the nested grouping of values can be observed. For that values are recursively merged into clusters of two, bases on a measure of similarity or distance, this is visualised by a vertical connection in the dendrogram (Rokach, L., 2010). The height of the connection between the two values represents the distance between them.

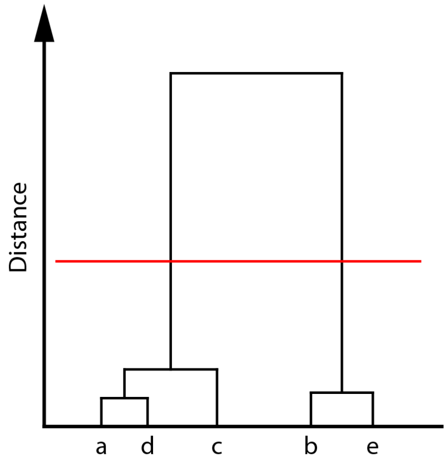

Figure 3: Exemplary Dendrogram showing multiple values connected. The height of the vertical lines represents the distance between to value or a cluster and a value. The red line represents a cut off limit where the dendrogram is split into multiple clusters.

In the first step of the clustering algorithm the pair of values with the closest distance is identified. These two values are then joined into a new cluster, that is represented by a new node connecting the two values in the dendrogram. The height of the new node is determined to be half the distance between the two values, make the new node equidistant to the two original values. As a result of the joining of the two values into a cluster a new set of distances between the new cluster and the remaining values needs to be calculated. This is done by a linkage function. There are multiple linkage functions known in the literature and many of them are available in the typical software packages used for hierarchical clustering. The linkage function D used in this work is the complete linkage function. As shown in Eq. (2) the equation updates the distances by computing the maximum distance of the individual values of the cluster.

$$D\big((a,d),c\big) = max(d(a,c),d(d,c)) \qquad \text{Eq. (2)}$$

After the distances have been calculated the algorithm repeats itself until only one cluster remains. The so gained dendrogram can then be used to identify the clusters of strongly correlated values. To get multiple clusters of strongly correlated pressure readings the dendrogram is partitioned into clusters. For that a cut-off limit has to be determined, that splits the dendrogram into multiple clusters. The values of each branch below the cut-off limit then form a cluster.

2.4. Structure identification and fault detection

Once the clusters of strongly correlated pressure readings have been determined by the hierarchical clustering algorithm, these clusters can be further investigated. Since the members of the clusters are known this information can be used to infer the layout of the supermarket from the data. Each cluster form a unique pressure level in the supermarket and all the members of the cluster are connected by a common suction line.

As the number of clusters should represents the number of pressure levels, a simple method of fault detection is to compare the number of identified clusters with the number of expected pressure levels. This method only requires the number of expected pressure levels to be known. For most Refrigeration systems this is known from the controller configuration. In addition to that, for most CO2-Supermarket systems it can be assumed that the number of pressure levels is equal to two, the MT and LT pressure level. If fewer or more clusters are identified this is a clear indication for a fault, warranting further investigation by a human operator. For this an additional time series plot of the evaporation pressures is created, wherein each identified cluster is plotted in one individual colour.

3. RESULTS AND DISCUSSION

The methodology is applied to a supermarket in Denmark which is equipped with a CO2 Supermarket refrigeration system, that has twelve individual evaporators. The layout of the supermarket is known. At first it is applied to data with no fault present, then two sets of data are analysed, where fault were present. In Figure 4 (left) the unordered correlation matrix of the evaporation pressures is shown. Using the hierarchical clustering method of section 2.3, the correlation matrix can be sorted by the resulting clusters, as shown in Figure 4 (right).

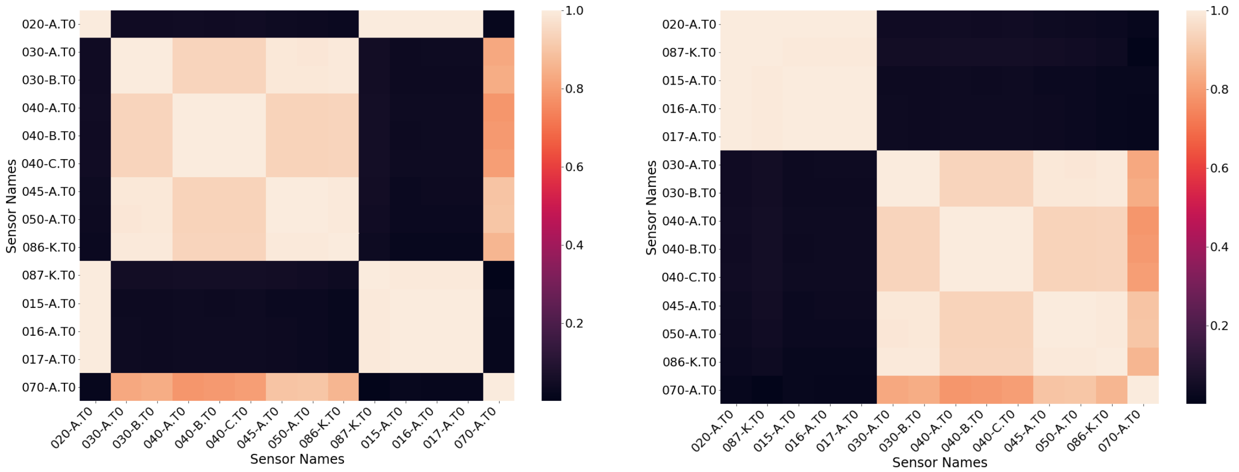

Figure 4: The unsorted correlation matrix of supermarket (left). The correlation matrix sorted by the identified clusters (right). The names on the axis are the names of the evaporators as well as the pressure sensors on the compressor rack.

For the clustering algorithm the cut-off limit used here was set at one third the total height of the dendrogram. Although there are some methods for automatically determining the cut-off limit, it was found in this work, that the most reliable way for achieving good splits was manually providing a fixed relative height. As the absolute values of the height in the dendrogram do not carry any significant value, it is better to use the fixed relative height for the cut-off. This also allows the relative height to be a user tuneable parameter to tune further analysis to his needs

In Figure 4 it can be seen, that two distinct clusters are present in the system. This result corresponds well to our expectations, based on the physical understanding of the system. In Figure 4 the names on the axis are the names of the evaporators in the monitoring system. In addition to that the names "086-K.T0" and "087-K.T0" are the pressure sensors on the compressor rack. A comparison of the two clusters with the real system layout showed the correctness of the results. Knowing this association of the connections between the individual evaporators and the compressor rack can enable advanced digital services, as for example the automatic creation of simulation models for digital twins.

The structure identification could also be replicated on two more supermarkets, where data was available and the layout was known. In these cases, the sampling rate for the data varied between one minute and fifteen minutes. The interval used for the evaluation was one week.

3.1. Example 1: Unexpected third pressure level

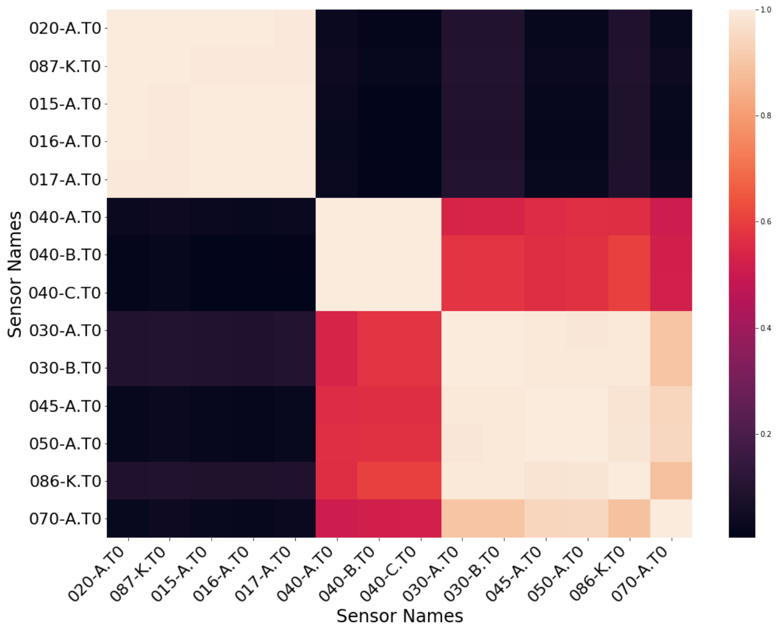

Figure 5: Pathological correlation matrix showing a third and unexpected pressure level

The first example of a fault in the supermarket system is shown in Figure 5, based on the correlation matrix visualised here three distinct clusters of pressure readings are present. A visual analysis of the evaporation temperatures as shown in Figure 6 seems to confirm the results, but with the knowledge, that in this supermarket system only two pressure levels should be present, this clearly is a fault in the system. Referring to the layout of the supermarket reveals, that these display cabinets should belong to the MT pressure level.

As far as it can be seen from the data, this error was present in the system from its first start after commissioning till remodelling of the supermarket. During this time the error did not cause any operational problems, that would have led to its discovery. Regardless of that, it still resulted in an unintended and potentially harmful operation of the display cabinets. As the pressure sensor is a necessary part in the calculation of the superheat at the outlet of the evaporators, when the controller sees an evaporation pressure lower than the actual evaporation pressure in the system, the calculated superheat will be larger than it actually is. Given a correct measurement of the refrigerant outlet temperature, the actual superheat the controller achieves is lower than the intended superheat. Depending on the absolute difference between measured and actual evaporation pressure, in the worst case this can lead to flooded operation. A closer evaluation of the valve opening degree of the expansion valve of one of the display cabinets shows a nearly full opening of the valve when the display cabinet is running. This suggests, the superheat controller is trying to lower, what it perceives to be, a high superheat, while in reality overfeeding the evaporator.

While in this case no adverse events, as a consequence of the fault, are known. In general, these silent faults can have severe consequences. One can easily imagine a scenario where the superheat controller overfeeding the evaporator results in one or more defect compressors and possibly a total loss of cooling.

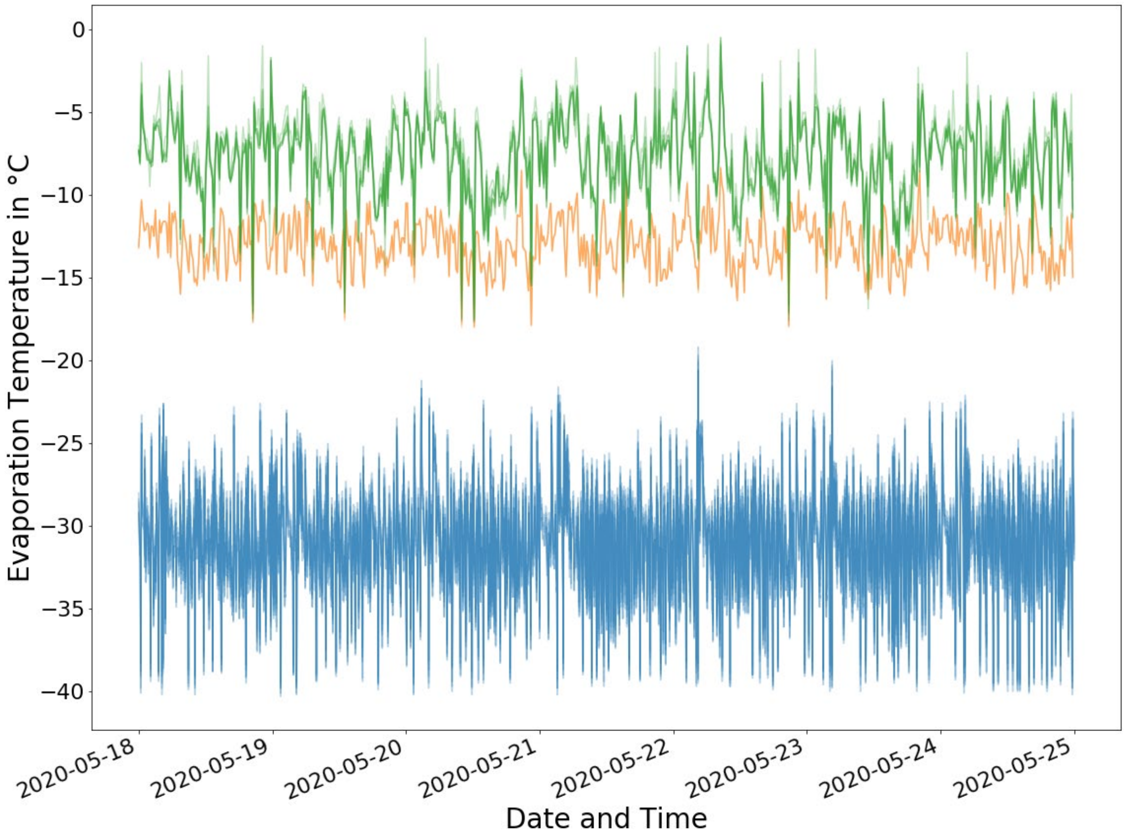

Figure 6: Evaporation temperature readings, coloured by the identified pressure levels. Clearly a supposedly third pressure level can be seen (orange). Based on the actual system layout these pressure readings should align with the MT pressure level (green)

3.2. Example 2: Sensor shows a constant value

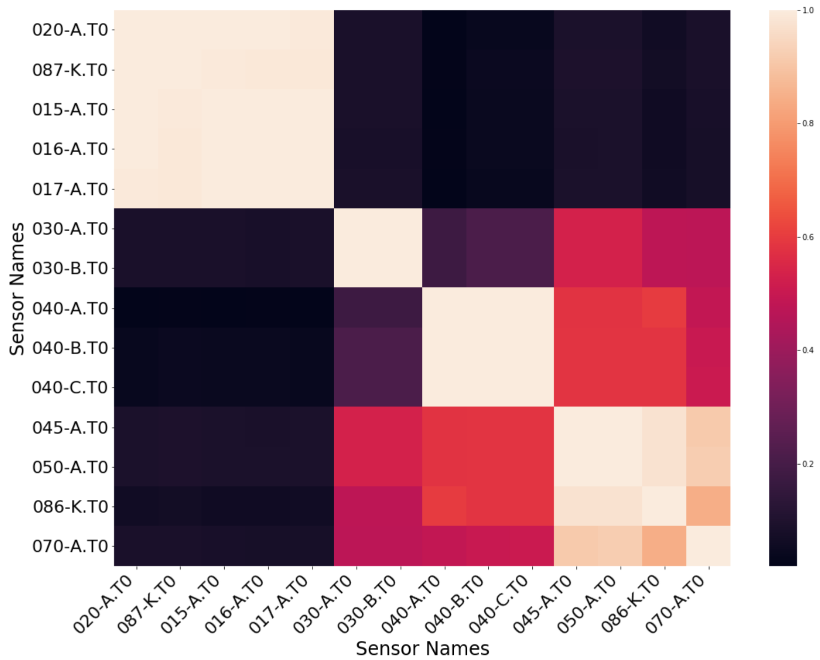

Figure 7: Pathological correlation matrix showing a addition forth pressure level, in addition to the already pathological third pressure level.

As a second example a fault was chosen, as shown in Figure 7. Here, in addition to the already known fault of a third cluster a fourth cluster appears in the correlation matrix, indicating another fault. In Figure 8 the evaporation temperatures are shown again and each detected pressure level is also coloured in a distinct colour. In this diagram now, there is a sensor which shows a near constant value that is also slightly higher than the average value of the other sensors. Again, with the knowledge of the actual supermarket layout, this sensor should belong to the MT pressure level. What makes this fault particularly interesting is the fact, that the pressure reading fails to a constant value, that lies within the expected range. In general controllers check for pressure readings outside the plausible range of values and give an alarm.

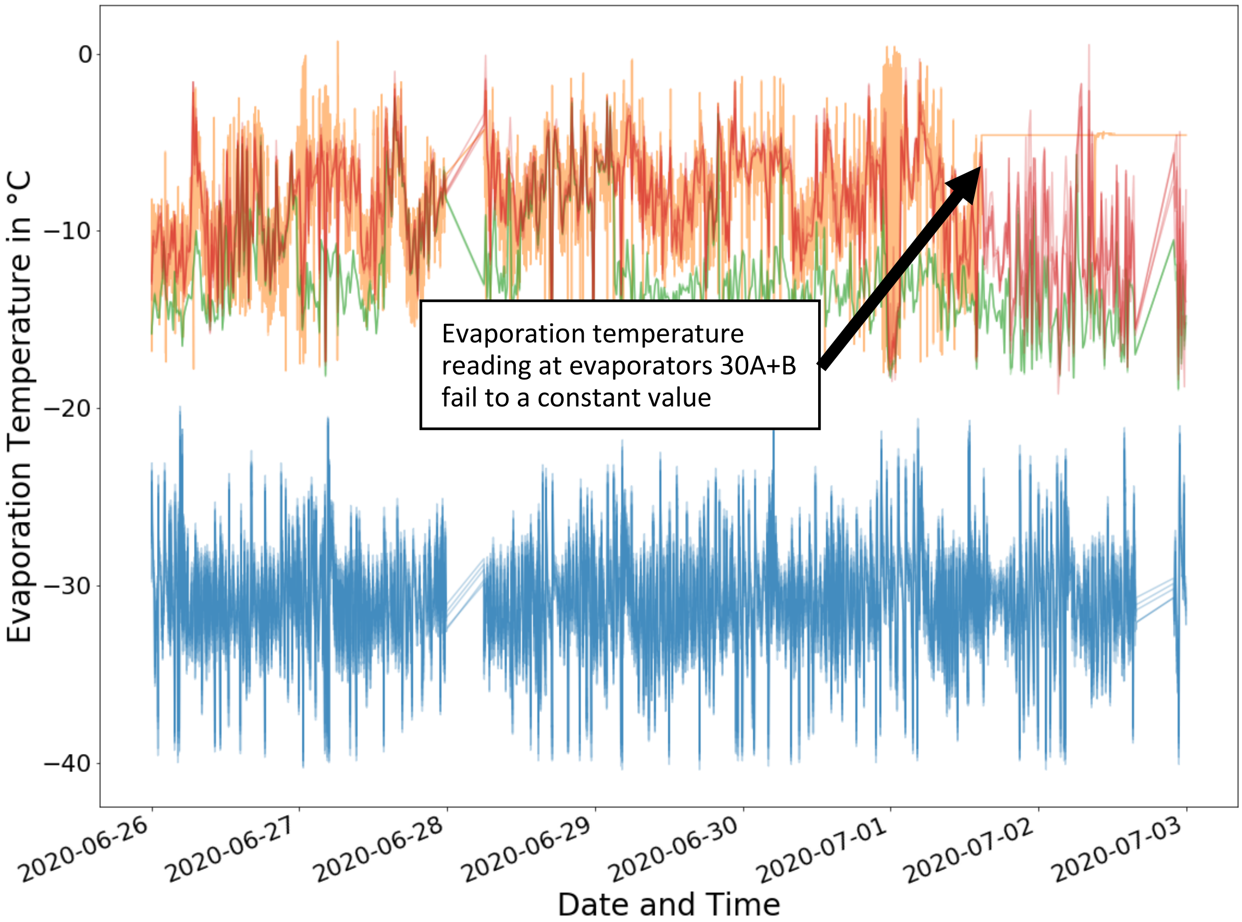

Figure 8: Evaporation temperature readings, coloured by the identified pressure levels. A the 1st of August one reading fails to constant value.

The operational consequences of the fault present an interesting case study on the effects of faults on complex and distributed systems, such as supermarkets. In a first step the local consequences of the fault need to be evaluated. The pressure reading failed to a plausible, yet slightly higher evaporation temperature value, than the actual evaporation temperature. As a consequence, the superheat calculated by the controller is smaller than the real superheat. To reach the required refrigerant outlet temperature for the superheat, under this condition, the controller needs to close the expansion valve further, compared to normal operation. As long as a temperature difference between the refrigerant at the outlet and the approaching air can be maintained, this poses no significant operational problem. Once the refrigerant at the outlet has the same temperature as the approaching air, no more heat can be transferred. If the controller then continues to close the expansion valve further, the reduced mass flow of refrigerant results in a reduced cooling capacity at the evaporator. If the fault only had local consequences, the reduced cooling capacity would ultimately result in a “high temperature” alarm and the defect sensor would most likely be found.

In this case the consequences of the fault where not only local. The supermarket in this example was running suction pressure optimization at the time of failure. This means that a higher-level control is implemented, that controls the evaporation pressure set-point. The intended idea is, to save energy by adapting the evaporation pressure of the refrigeration system to the systems actual load. For that the controller looks at all evaporators connected to the system and determines the “most loaded case” (MLC), in effect the MLC is the evaporator on the system that is struggling the most to maintain temperature. Based on the MLC the set-point for the suction pressure is increased or decreased (Danfoss, 2021). In this example, as a result of the reduced cooling capacity, it follows, that the evaporator becomes the “most loaded case”. As a consequence, the higher-order controller decreases the suction pressure set-point, in what it believes to be an increased load on the system. Figure 9 shows the evaporation temperatures of evaporators 30A and 30B, as well as the set-point for the evaporation temperature and the actual evaporation temperature at the compressor rack, before and after the fault occurred. It can be seen, that the evaporation temperature decreases drastically, the moment the fault occurs.

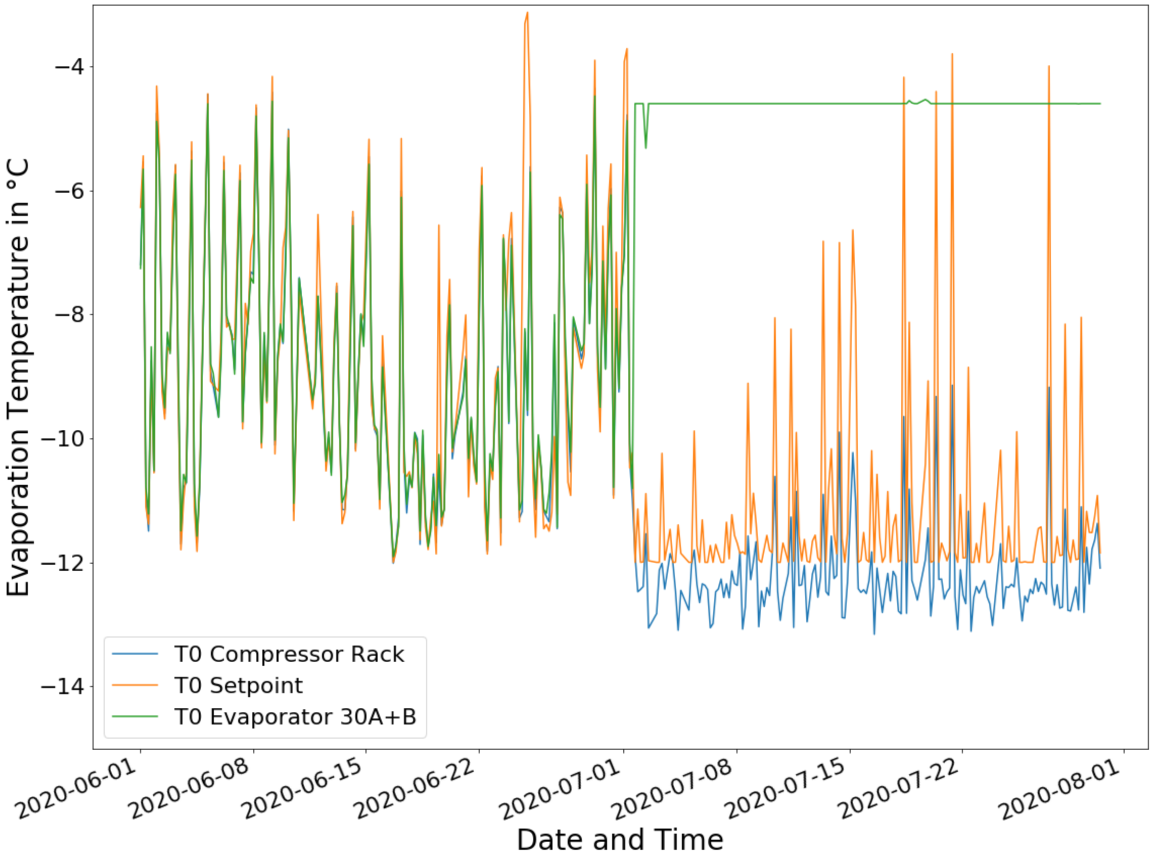

Figure 9: Evaporation temperature reading at evaporator 30A+B, as well as the compressor rack. Additionally, the setpoint for the evaporation temperature of the compressor rack. Hourly average values shown.

Because the suction pressure optimization lowered the evaporation pressure, the fault of the pressure reading is compensated and the temperature could be maintained. Because the temperature was maintained, this fault remained silent and no directly adverse event happened. But as an indirect consequence of this fault, an increased energy demand for the system occurred.

CONCLUSIONS

A method for structure identification and fault detection using the measured evaporation pressures in a supermarket refrigeration system was presented. The method is based on the analysis of the correlation of the evaporation pressures and uses hierarchical clustering to identify the individual pressure levels. It was shown, that using this method the components belonging to a pressure level can be identified. Using a priori knowledge of the number of pressure levels in the system this method can also be used to identify faults in the refrigeration system. Two examples of fault identified with this method were shown. Both of them were difficult to detect by conventional methods, as neither of them resulted in difficulties maintaining temperature. Despite that, the first fault posed a potential risk to the compressors and the second fault resulted in increased energy consumption. Both would have required extensive manual work, evaluating the monitoring data, to identify these faults, emphasising the need for automated fault detection methods.

ACKNOWLEDGEMENTS

This work was funded by the EUDP - Energy Technology Development and Demonstration Program under Grand EUDP 64019-0570.

We greatly appreciate AK Centralen (https://ak-centralen.dk/) for providing access to the data used in this publication.

NOMENCLATURE

D	Linkage function
d	Distance measure

REFERENCES

Danfoss, 2018. User Guide Controller for evaporator control AK-CC 750A. Danfoss A/S, Nordborg, 105.

Danfoss, 2021. User Guide AK-System Manager AK-SM 800A series. Danfoss A/S, Nordborg, 102

European Commission, 2014. The economic impact of modern retail on choice and innovation in the EU food sector. Publications Office of the European Union, Luxembourg, 452.

Rokach, L., 2010. A survey of Clustering Algorithms. In: Maimon, O., Rokach, L., Data Mining and Knowledge Discovery Handbook, Springer, New York, 269-299.

Zeidler, E., 2013. Springer-Taschenbuch der Mathematik. Springer Fachmedien Wiesbaden, Wiesbaden, 1324 pp.