A Commissioning-Oriented Fault Detection Framework for Building Heating Systems Using SARIMAX Models

Parantapa Sawant[1], Ralph Eismann[1]

[1] Institute for Sustainability in Energy and Construction (INEB),
University of Applied Sciences Northwest Switzerland (FHNW)
parantapa.sawant@fhnw.ch

Abstract

A scalable and rapidly deployable fault detection framework for building heating systems is presented. Unlike existing data-intensive machine learning approaches, a SARIMAX-based concept was implemented to address challenges with limited data availability after commissioning of the plant. The effectiveness of this framework is demonstrated on real-world data from multiple solar thermal systems, indicating potential for extensive field tests and applications for broader systems, including heat pumps and district heating.

Keywords: Building Technologies; Data-Driven Fault Detection; SARIMAX.

1 Introduction

From 2019 to 2023, LoRaWAN-based temperature sensors were installed to monitor the outlet of approximately 450 building solar thermal systems [1]. We developed a rule-based algorithm (RBA) for fault detection and diagnosis (FDD) by leveraging extensive operational expertise and achieved around 95% accuracy. However, the RBA faced scalability issues due to the variations in plant characteristics, dependent on type of installation or control strategies amongst other factors, that cannot be captured with just one sensor per plant. While it meets practical requirements, our current goal is to develop a complementary data-driven algorithm capable of rapid fault detection (FD) across various installations, and potentially extendable to heat pumps and district heating systems. This work presents the results of the first proof-of-concept. The ability of machine learning (ML) algorithms for FDD in building HVAC systems especially solar thermal systems is well documented [2–9]. These studies employed extensive process-history data from multiple sensors, simulation data, and used complex models such as random-forest-regression to detect numerous faults and in certain cases, diagnose them. We also attempted to train ML models using results of the RBA but were hindered by a lack of labeled data representing ground truth and the intrinsic limitations of training ML models to replicate the RBA [10]. Instead, we propose a novel approach utilizing a time series forecasting model that strikes a balance between scalability and minimal data available post-commissioning, facilitating preliminary FD using only the single data-point per installation.

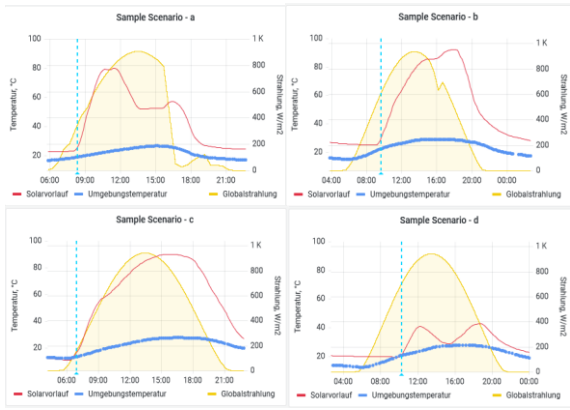

Fig. 1. Examples for different scenarios that may occur in a solar thermal plant.[1]

Fig. 1 shows the solar output temperature for four different installations under similar environmental conditions. For example, scenario a and scenario b: A monotonic, concave temperature increase over a period until a rapid temperature change is observed due to cloud cover or peak load, etc. Scenario c: A completely monotonic, concave temperature curve until sunset or lower solar radiation. Scenario d: A special case where none of the above occurs, although there is an initial temperature rise and a fault in the system could be the cause. These are just a few examples among many others, including night cooling or stagnation in solar thermal systems, for which the reader is referred to previous work by the authors [1]. Different types of systems will behave differently due to their design and control measures for similar ambient temperatures and global irradiance. A single model trained on all data will be limited in predicting failures for all plants. We want to explore an approach that uses minimal and readily available historical data from a plant to be implemented for that plant and easily integrated into a pipeline for widespread implementation. The concept of implementing this pipeline is also discussed in this paper.

2 Methodology

The data was collected from solar thermal systems at 10-minute intervals, and for this project we have one to two years of data per system. However, since future installations may not have extensive historical data, the focus is on developing a model that requires minimal data for fitting. So, for the exploratory analysis and visual analysis of the time series we used the entire data set of a plant and for the model training and fitting phase we assume 3 past days are enough for predicting the target day. This assumption was empirically evaluated by comparing training models using past 1 and 2 days and the results are discussed later in this paper. The data handling tasks were performed using *Python-Pandas v2.1.1*, the model fitting and diagnosis was performed using the *statsmodels v0.14.0* [11] and the *pmdarima v2.0.4* [12] package on a laptop equipped with an 11th Gen Intel(R) Core(TM) i9-11950H @ 2.60GHz and 32GB Ram.

Duplicate timestamps were identified and removed, with only the first occurrence retained to maintain temporal accuracy. Missing values were addressed using linear interpolation for gaps up to four hours, reducing the percentage of missing data from 0.30% to 0.13%, thus preparing the dataset for subsequent analysis.

[1] Solarvorlauf/Collectorvorlauf = Collector outlet temperature, Umgebungstemperatur = Ambient temperature, Globalstrahlung = Global irradiation

2.1 Exploratory Data Analysis (EDA)

Each dataset comprises four key variables: collector outlet temperature (in Celsius), representing the temperature of the fluid exiting the solar thermal collectors; ambient temperature (in Celsius), which measures the surrounding environmental temperature; cloud cover (in percentage), indicating the percentage of cloud coverage, with 100% representing fully cloudy conditions; and global irradiation (in Watts per square meter), measuring the total solar radiation received per square meter. A statistical analysis was conducted on these variables, calculating their mean, minimum, and maximum values, and producing box plot visualizations for the purpose of assessing data quality, identifying outliers, and evaluating the spread of values. It was found that all sensor readings fell within the expected operational ranges for Switzerland, thereby confirming the validity of the dataset for further analysis. The correlation between variables was assessed using *Spearman's rank correlation coefficient*, as shown in **Table 1**. A strong positive linear correlation was detected between collector outlet temperature and both ambient temperature and global irradiation. The effect of cloud cover was deemed superfluous in the context of fitting, given that its influence is presumed to be encapsulated in the global irradiation effect. In consideration of the considerable discrepancies in the magnitude of these variables, a scaling process was undertaken to ensure a mean of 0 and a standard deviation of 1.

Table 1. Correlation between collector output and other variables

	Dependant ↓ / Independent →	**Cloud Cover**	**Ambient Temp.**	**Global Irradiation**
Entire data set	Collector outlet temp.	-0.21	0.78	0.82
3 days subset	Collector outlet temp.	-0.21	0.92	0.87

2.2 Model Training

Prior to establishing the production pipeline, the *Box-Jenkins methodology* was utilized to ascertain, quantify and diagnose the model. The identification of the three-day time series was conducted through the application of additive decomposition, with the resulting data for a single set illustrated in the **Fig. 2** below. Despite the absence of an overall trend, which was anticipated for solar thermal collectors, pronounced daily seasonality was discernible.

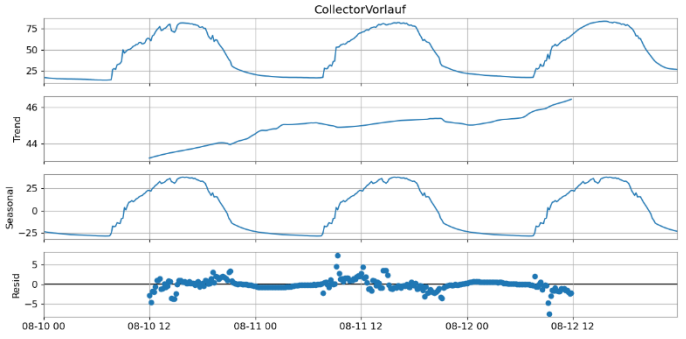

Fig. 2. Additive decomposition of the 3-day training data set

It was determined that a *Seasonal Autoregressive Integrated Moving Average with Exogenous Variables* (SARIMAX) model would be the most appropriate means of capturing the identified seasonal patterns and relationships between variables. An *Augmented Dickey-Fuller* (ADF) test was performed to check for stationarity. The results are shown in **Table 2**, with a single seasonal

differencing applied. Significant reduction in the ADF-statistic with the seasonal differencing informed the decision to use single differencing for the seasonality.

Table 2. Results of an Augmented Dickey-Fuller test for stationarity

Series	ADF-statistic	p-value
Original	-3.8	0.002
First Difference	-2.9	0.04
Seasonal Difference	-5.5	1.6e-06

Model parameters were further refined using the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, which informed the selection of a first order autoregressive (AR) model. The ACF and PACF plots for the training data, plotted for 5-hour lags (equivalent to 30 lags for 10 minute resolution data), show a clear seasonal pattern. This seasonality is evident from the cyclical nature of the ACF plot, while the PACF plot shows a significant cut-off after the first lag.

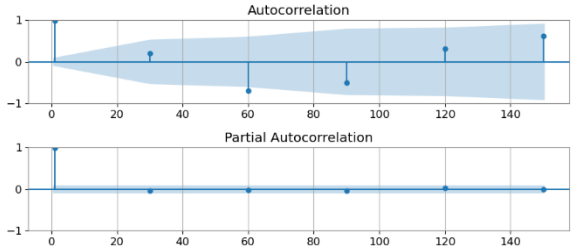

Fig. 3. ACF and PACF plots for training data set

These initial estimates were used as a basis for parameter tuning by grid search, using the *pmdarima* package to automate the selection of an optimal SARIMAX model based on minimizing the *Akaike Information Criterion* (AIC). The refined SARIMAX model was then applied to predict the outlet temperature for a given target day, using solar radiation and ambient temperature forecasts as exogenous regressors. The model's prediction was evaluated against actual measurements using the *Root Mean Squared Error* (RMSE) criterion, and the target day was classified as either a 'fault day' (F-Day) or a 'no fault day' (NF-Day) based on predefined accuracy thresholds (10 K in our case). This process was integrated into a machine learning (ML) pipeline designed to run once a day for each plant. For each iteration, the last three NF-days were used as the training data set for the SARIMAX model, ensuring continuous adaptation to changing conditions.

The methodology described above is illustrated in the framework shown in FIGURE. In this framework, the user first defines a training dataset consisting of as few as three consecutive NF-days. This data set is used to fit the SARIMAX model, which generates predictions for the following day. If the predicted and observed values for the target day are sufficiently close, the day is labelled an NF-day, otherwise it is labelled an F-day. The training dataset is then updated with the latest NF-days to ensure that the SARIMAX model is always trained on the latest operational data. In practical applications, this loop is run once a day for each sensor, and the initial training data only needs to be defined once for each sensor.

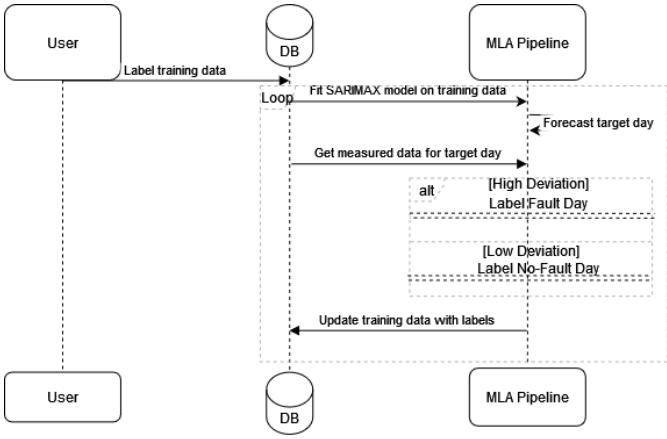

Fig. 4. Exemplary framework for implementing the MLA pipeline in the production system of industry partners

Results of this approach are shown in the next section.

3 Results

To facilitate concise presentation, the results presented here are from a single plant that exhibited the most stable performance, characterized by minimal NF-days. This allows effective model validation. A manual review of all historical data to identify NF-days would be impractical, and the limitations of the ground truth are discussed in the Introduction section. The results of the various tests, in which the dataset lengths were either 2 or 3 days, demonstrate the performance across different types of days selected to represent variations in global irradiation and ambient temperature (see Fig. 5). The analysis demonstrates that the automated fitting process identifies distinct optimal parameters for the SARIMAX model contingent on the length of the dataset. It is notable that while the fitting time of the grid search is shorter with a 2-day dataset, this is accompanied by a reduction in RMSE accuracy. In contrast, the 3-day tests, while taking longer to fit, tend to yield more reliable predictions, indicating a trade-off between model complexity, accuracy, and computational efficiency. These findings highlight the necessity of meticulously determining the optimal dataset length to achieve a balance between the necessity for a robust model performance and the practical considerations of fitting time and computational resources.

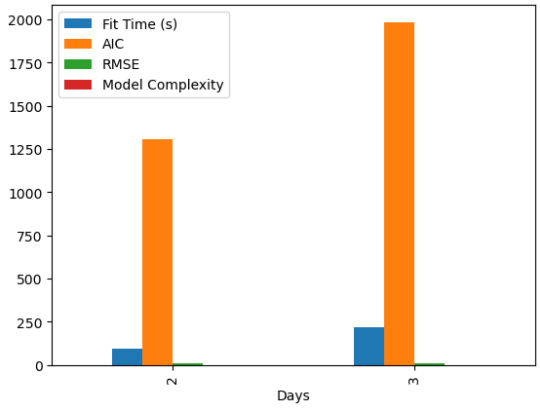

Fig. 5 Summary of hyperparameter search and results for various test data sets

Fig. 6 and **Fig. 7** illustrate the iterative deployment of the ML pipeline designed to analyze three target days, utilizing temperature data from their preceding three NF-Days for training. The autofitting of the SARIMAX model is employed in this context. In **Fig. 6**, all three forecasted days are classified as NF-Days and subsequently appended to the NF-Day list, which is then utilized for further model fitting and refinement. Conversely, in **Fig. 7**, the algorithm effectively identifies F-Days, which are characterized by significant deviations from the expected behavioral patterns of the system. Such deviations may arise due to various factors, including faulty hydraulic configurations, suboptimal controller settings, inefficient pump operations, or adverse environmental conditions. While diagnosing these anomalies was beyond the scope of the current algorithm, it is critical to acknowledge that some deviations could be attributed to rapidly fluctuating global irradiation, often influenced by cloud cover dynamics. The complexity of SARIMAX model fitting underscores the necessity for rigorous evaluation and validation of the model's performance. Selecting appropriate training days is a critical step that requires expert judgment to ensure the representativeness and relevance of the data utilized. The expert selection process helps mitigate potential biases and enhances the model's capacity to generalize across different operational conditions. Future enhancements could involve implementing the algorithm primarily on days characterized by minimal cloud cover to reduce variability in the data. Additionally, adjusting the weights of exogenous variables within the SARIMAX model fitting could further improve the model's responsiveness to external influences. Furthermore, exploring alternative algorithms that better capture the effects of rapidly changing global irradiation remains a valuable avenue for future research, although it lies outside the current study's objectives.

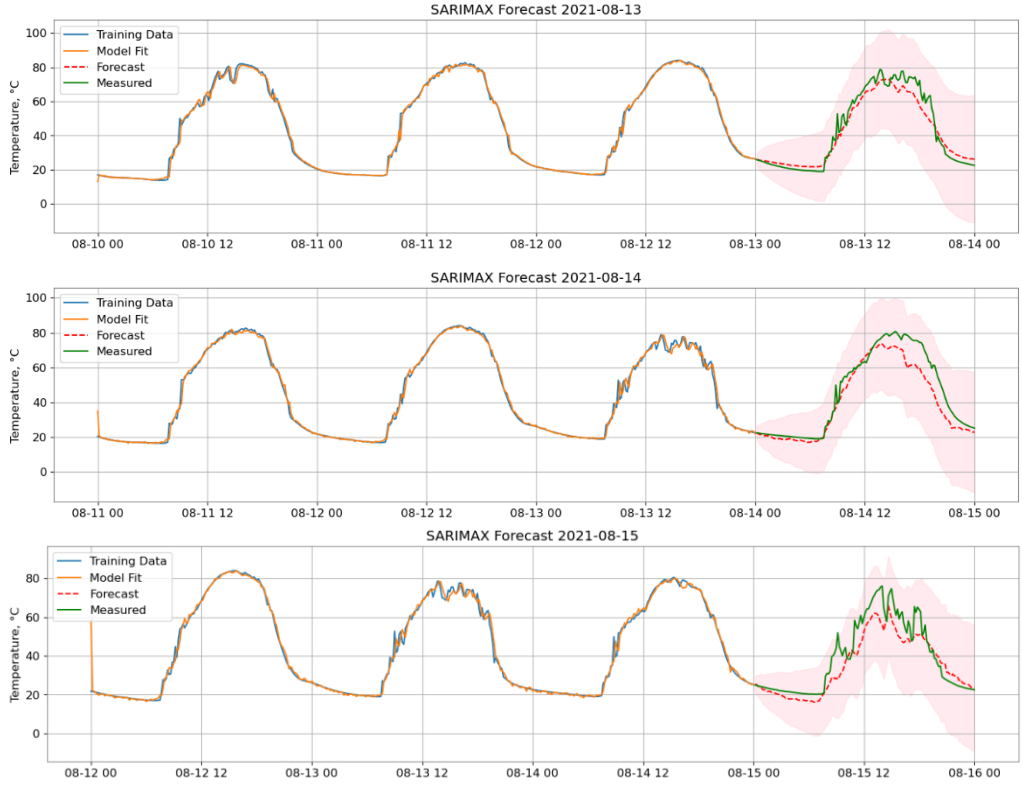

Fig. 6. Results for three iterations of the ML-Pipeline on non-cloudy days in August for plant ST4048

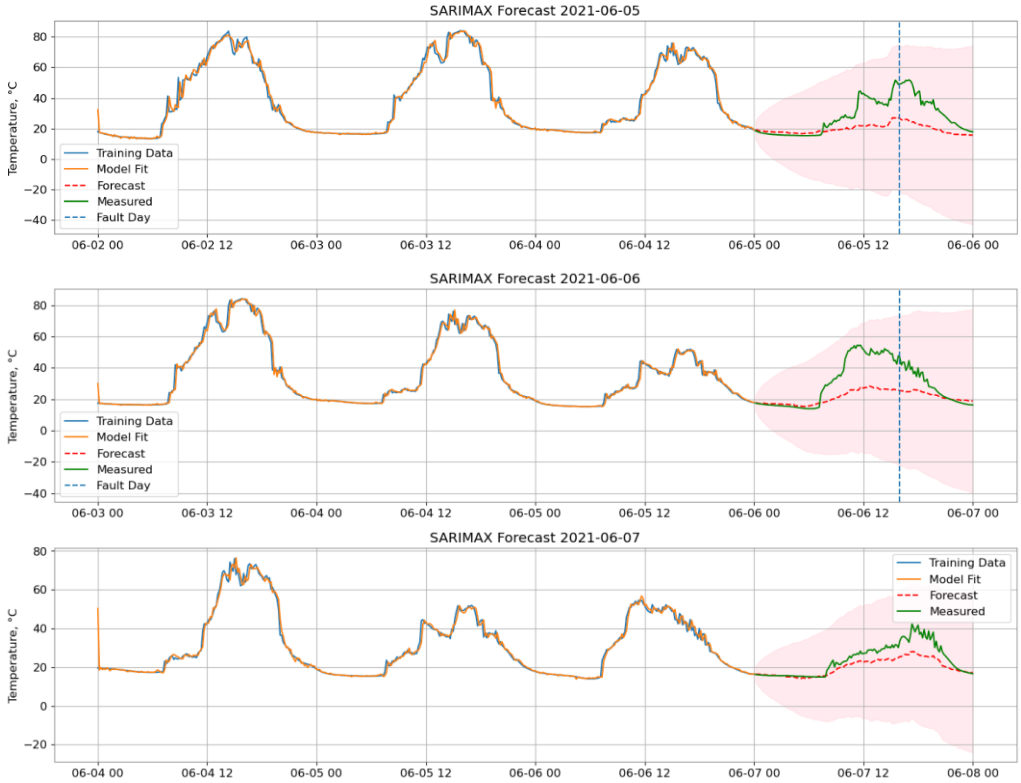

Fig. 7. Results for three iterations of the ML-Pipeline on cloudy days in June for plant ST4048

For the quantification of the results, a confusion matrix was constructed for various tests, utilizing subsets of the entire data that predominantly contained NF-days, as identified by field experts. This approach was necessary since ground truth verification could not be conducted otherwise, and the datasets lacked a specific variable to decisively classify a day as either F-day or NF-day.

The results, summarized below, reveal that the autotuned SARIMAX model underperformed when compared to the manually tuned model, which used fixed parameters across all iterations. The autotuned model tended to falsely identify more F-days, indicating higher rates of false positives, while the manually tuned model with fixed order exhibited better consistency and fewer misclassifications. This highlights the challenges in automating model tuning and the importance of expert input, particularly when ground truth data is not readily available for validation.

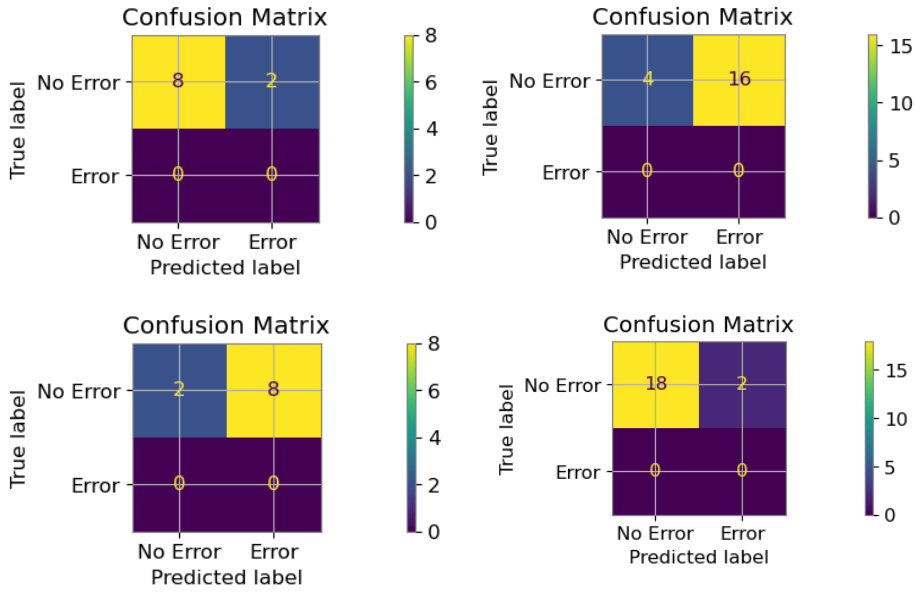

Fig. 8. Confusion matrix results for various tests. Clockwise: (a) SARIMAX(1,1,1)(1,1,0)[144] for 10 cloudy days, (b) Autotuned for 20 mixed days, (c) SARIMAX(1,1,1)(1,1,0)[144] for 20 cloudy days , (d) Autotuning for 10 cloudy days

4 Conclusion

This study has demonstrated the effectiveness of a SARIMAX-based framework for fault detection in solar thermal systems that uses minimal historical data while still providing reliable results. By incorporating solar irradiance and ambient temperature as exogenous variables, the proposed model successfully adapts to dynamic operating conditions with only three fault-free days required for training. This makes the solution scalable and applicable to installations where limited data is available after commissioning.

Automated model updating through a machine learning pipeline further enhances the system's ability to adapt to changing conditions with minimal recalibration. This lays the groundwork for practical applications not only in solar thermal systems, but also in broader systems such as heat pumps and district heating, which can benefit from this streamlined approach.

Future work should focus on extending the framework to other heating systems, including more diverse test environments, refining the accuracy of error detection, and incorporating additional exogenous factors such as collector orientation and refined weather data filters. The results of this study highlight the potential for a scalable, low-data solution to monitor and optimize building heating systems, contributing to their overall efficiency and reliability.

5 Acknowledgements

We would like to express our gratitude to *Verein Energie Zukunft Schweiz* for providing the data used in this study, which was collected as part of the LoCoSol+ project.

References

[1] P. Sawant, B. Sintzel, R. Eismann, J.W. Hofmann, B. Sitzmann, Low Cost Monitoring thermischer Solaranlagen mit IoT-Sensor und maschinellem Lernen, (n.d.). https://www.aramis.admin.ch/Texte/?ProjectID=49446 (accessed July 11, 2024).

[2] G. Faure, M. Vallée, C. Paulus, T.Q. Tran, Fault detection and diagnosis for large solar thermal systems: A review of fault types and applicable methods, Solar Energy 197 (2020) 472–484. https://doi.org/10.1016/j.solener.2020.01.027.

[3] L. Feierl, V. Unterberger, C. Rossi, B. Gerardts, M. Gaetani, Fault detective: Automatic fault-detection for solar thermal systems based on artificial intelligence, Solar Energy Advances 3 (2023) 100033. https://doi.org/10.1016/j.seja.2023.100033.

[4] Z. Liu, Y. Liu, D. Zhang, B. Cai, C. Zheng, Fault diagnosis for a solar assisted heat pump system under incomplete data and expert knowledge, Energy 87 (2015) 41–48. https://doi.org/10.1016/j.energy.2015.04.090.

[5] M.S. Mirnaghi, F. Haghighat, Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review, Energy and Buildings 229 (2020) 110492. https://doi.org/10.1016/j.enbuild.2020.110492.

[6] H. He, T.P. Caudell, D.F. Menicucci, A.A. Mammoli, Application of Adaptive Resonance Theory neural networks to monitor solar hot water systems and detect existing or developing faults, Solar Energy 86 (2012) 2318–2333. https://doi.org/10.1016/j.solener.2012.05.015.

[7] R. Kicsiny, Z. Varga, Real-time nonlinear global state observer design for solar heating systems, Nonlinear Analysis: Real World Applications 14 (2013) 1247–1264. https://doi.org/10.1016/j.nonrwa.2012.09.017.

[8] C. De Keizer, S. Kuethe, U. Jordan, K. Vajen, Simulation-based long-term fault detection for solar thermal systems, Solar Energy 93 (2013) 109–120. https://doi.org/10.1016/j.solener.2013.03.023.

[9] S. Jiang, M. Lian, C. Lu, S. Ruan, Z. Wang, B. Chen, SVM-DS fusion based soft fault detection and diagnosis in solar water heaters, Energy Exploration & Exploitation 37 (2019) 1125–1146. https://doi.org/10.1177/0144598718816604.

[10] J.W. Hofmann, B. Sitzmann, J. Dickinson, D. Kunz, R. Eismann, Use of machine-learning for monitoring solar thermal plants, Journal of Physics: Conference Series 2042 (2021) 012007. https://doi.org/10.1088/1742-6596/2042/1/012007.

[11] Josef Perktold, others, statsmodels/statsmodels: Release 0.14.1, (2023). https://doi.org/10.5281/ZENODO.593847.

[12] T.G. Smith, others, pmdarima: ARIMA estimators for Python, (2017). http://www.alkaline-ml.com/pmdarima.