Received: January 1, 2024. Revised: February 10, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Optimizing Air Handling Unit Induction Motor Frequency Predictions: Evaluating and Advancing Forecasting Techniques with a Modified Chen's Fuzzy Time Series Model

Yaddarabullah[1]* Yoga Alviando[1] Dewi Lestari[1] Aedah Binti Abd Rahman[2] Amna Saad[3]

[1]*Department of Informatics, Universitas Trilogi, Jakarta, Indonesia*
[2]*Schools of Science and Technology, Asia e University, Selangor, Malaysia*
[3]*Malaysian Institute of Information Technology, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia*
* Corresponding author's Email: yaddarabullah@trilogi.ac.id

Abstract: The purpose of this study is to introduce a novel approach to predict induction motor frequency adjustments in Air Handling Units (AHU). This is essential as traditional methods have frequently been unable to effectively address the complex seasonal and stochastic fluctuations that are inherent in these environments. To overcome this challenge, the research focuses on utilizing experimental Chen's Fuzzy Time Series model, specifically designed to incorporate temporal and seasonal patterns into the predictive analysis. Various predictive models, including Seasonal Autoregressive Integrated Moving Average (SARIMA), Holt-Winters Exponential Smoothing (HWES), neural network (NN) based ensemble model, hybrid of artificial neural network (ANN) and SARIMA, and Seasonal Autoregressive Integrated Moving Average with Exogenous factors (SARIMAX), are compared to determine the most effective model in optimizing AHU induction motor frequency. Results indicate that the modified Chen's Fuzzy Time Series model demonstrated high efficacy with an R-squared value of 0.9945 in a one-hour time interval in the seasonal pattern, showing an almost perfect fit between predicted outcomes and actual data compared to other prediction models. Furthermore, the modified Chen's model achieved a Mean Absolute Percentage Error (MAPE) of 2.41% and a Root Mean Square Error (RMSE) of 0.72, significantly outperforming other models in predictive accuracy and reliability. The modified Chen's model showed an efficiency improvement of 86% in MAPE and 87% in RMSE compared to other prediction models.

Keywords: Air handling unit, Induction motor frequency predictions, Chen's fuzzy time series, Seasonal variations, Time interval, Stochastic variability, Forecasting optimization, Predictive modelling.

1. Introduction

Heating, Ventilation, and Air Conditioning (HVAC) systems are essential components of contemporary buildings, designed to regulate indoor environments for the purpose of ensuring comfort, air quality, and appropriate temperature [1]. The heating element often utilizes boilers or furnaces, particularly in colder climates, to provide warmth. Ventilation is crucial for replacing or exchanging air within a space to control temperature, replenish oxygen, remove moisture, smoke, heat, dust, airborne bacteria, and carbon dioxide, and to replenish oxygen. Air conditioning, which is commonly combined with ventilation systems, removes heat and moisture from the interior of the building to maintain a comfortable and healthy environment. The functioning of buildings is considerably affected by Heating, Ventilation, and Air Conditioning (HVAC) systems, particularly in locations where human activity is continuous, such as offices, schools, and institutional spaces [2]. The importance of a dependable HVAC system cannot be overstated. Not only does it ensure effective air circulation, but it also plays a crucial role in maintaining a comfortable ambient temperature and

humidity level within a room [3]. A widely utilized approach in diverse environments is the implementation of centralized HVAC systems, which efficiently and methodically circulate air throughout multiple rooms.

The Air Handling Unit (AHU) is an essential component of the HVAC system, as it is responsible for conditioning and distributing air throughout various areas. The AHU comprises several key elements, including an induction motor, filter, heating and cooling coils, dampers, and fans, which all work together to ensure optimal indoor air quality and temperature. Specifically, the induction motor drives the fan, which circulates and conditions the air [4]. The utilization of inverters enables the adjustable-speed control of induction motors employed in AHUs, thereby facilitating adaptability to fluctuating airflow or load conditions. This level of control is essential in ensuring consistent temperature and humidity levels, which ultimately contribute to the comfort and productivity of occupants [5]. The optimization of the frequency of the AHU induction motor is not only critical for enhancing energy efficiency, but also for ensuring effective environmental control. By adjusting the motor frequency in response to fluctuating air circulation needs, it is possible to optimize power consumption, thereby promoting a more energy-efficient and cost-effective operational approach [6]. It is essential to precisely regulate the airflow through the AHU induction motor frequency in order to maintain optimal temperature and humidity levels [7]. This helps prevent issues such as mold growth and ensures a healthy indoor environment. [8].

The importance of the AHU induction motor frequency is further underscored when considering the impact of occupant behaviour and seasonal fluctuations [9]. In commercial and academic buildings, the intensity of occupant activities and their patterns can vary significantly with the seasons, influencing the indoor environmental needs. In commercial buildings, for example, occupancy levels in spaces such as office areas, meeting rooms, and retail spaces fluctuate with business cycles, holidays, and seasonal sales periods [10]. In academic settings, the utilization of classrooms, lecture halls, and libraries changes with the academic calendar, peaking during active semesters and decreasing during breaks [11]. These variations necessitate the dynamic adjustment of the AHU induction motor frequency to cater to the evolving demands. Critical to this process is the selection of appropriate time intervals for regulating the AHU induction motor. Efficiently meeting the building occupants' varying needs requires a dynamic strategy that adapts to daily and seasonal changes in occupancy and environmental conditions [12]. For instance, in commercial buildings during typical workday hours, more frequent adjustments are necessary to manage peak occupancy levels [10].

Previous studies in the field of HVAC systems predominantly focused on conventional methods for controlling AHU (Air Handling Unit) induction motors. These methods primarily relied on reactive, scheduled, and predictive control. The reactive control approach responded directly to changes in environmental conditions as they occurred [13,14], while scheduled control operated based on predetermined timeframes, typically aligning with expected usage patterns of the building [15]. Predictive control utilizes weather forecasts to inform control decisions, while reactive controls respond after environmental changes, causing delays. Static schedules overlook seasonal occupancy variations, resulting in inefficient HVAC resource use. Predictive models offer a proactive approach, especially for addressing seasonal variations in HVAC control.

Jetcheva et al. [16] conducted research aimed at optimizing HVAC control systems within commercial and industrial settings, with a particular emphasis on accounting for seasonal variations. Their approach involved developing a predictive based on neural network-based ensemble model. To evaluate the efficacy of this model, they compared it with several established predictive methodologies, including Seasonal Autoregressive Integrated Moving Average (SARIMA), Holt-Winters, Random Forest, and Multiple Regression. The critical input data for their model comprised indoor climatic variables, namely temperature and humidity, ensuring a focused and relevant analysis within the context of HVAC control.

Liu et al. [17] developed a prediction model for HVAC system control, integrating techniques from Support Vector Machine (SVM), Artificial Neural Network (ANN), and Seasonal Autoregressive Integrated Moving Average (SARIMA). Their approach was particularly designed to account for seasonal variations. The experimental results from their study indicated that this hybrid HVAC control based prediction model effectively enhances prediction accuracy, outperforming the results achievable by each of the individual models when used in isolation.

Fathollahzadeh [18] developed a time-series prediction model for AHU induction control, utilizing the Seasonal Autoregressive Integrated Moving Average with Exogenous factors

(SARIMAX). This model, designed to consider seasonal variations, incorporates input data such as weather conditions, indoor temperature, humidity, and chilled water flow rate. The integration of these diverse data points allows for a comprehensive approach to predict AHU control needs.

Previous research highlights SARIMA and Holt-Winters as baseline methods for developing predictive models for AHU induction motors. Their effectiveness in capturing seasonal variations makes them preferred choices in this domain [19]. Moreover, state-of-the-art methods such as neural network based ensemble model [16], a hybrid of ANN and SARIMA [17], and SARIMAX [18] are also utilized for predicting AHU induction motor control. An additional advantage of these statistical methods is their efficiency in environments with limited data availability, as they do not require extensive datasets to perform effectively [20]. However, a critical aspect that previous studies have often overlooked is the selection of appropriate time intervals in these models. The correct determination of time intervals is crucial, as inaccuracies can lead to missed detection of important seasonal patterns, resulting in incorrect predictions [21]. Incorrect time intervals may cause these models to either ignore essential seasonal highs and lows or merge them with other data variations, thereby masking the true effects of seasonality [22]. Consequently, a significant aspect of advancing HVAC control technology lies in accurately identifying and applying these time intervals to ensure that seasonal variations are precisely captured and responded to.

The present study endeavors to enhance the precision of predicting AHU induction motor frequencies, specifically focusing on data collected from the Universitas Trilogi library during a semester marked by prominent seasonal fluctuations. To meticulously discern these seasonal variations, we meticulously applied the Friedman Test and the Entropy method to pinpoint the most suitable time intervals for our analytical pursuits. The study meticulously refined and experimented with Chen's Fuzzy Time Series model, aiming to surpass established seasonal prediction models such as SARIMA, Holt-Winters, SARIMAX, neural network-based ensemble models, and hybrid models of ANN and SARIMA. Our overarching objective was to elevate existing models and devise a more precise prediction methodology.

The primary thrust behind refining Chen's Fuzzy Time Series model centered on optimizing the defuzzification process to more accurately capture seasonal variations. Our methodology involved a systematic approach where we iteratively fine-tuned parameters and algorithms within the Chen's Fuzzy Time Series. This iterative refinement process aimed to enhance the model's sensitivity to subtle seasonal nuances, thus augmenting its predictive accuracy. To comprehensively evaluate the performance of these predictive models, we employed a suite of statistical metrics including Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared values. These metrics formed the cornerstone of our analytical framework, providing a robust means to assess each model's predictive efficacy concerning AHU induction motor frequencies. By meticulously scrutinizing the predictions against actual data, we gauged the models' ability to capture the intricate dynamics of AHU induction motor frequencies across various time intervals and seasonal patterns.

The organization of this research paper adheres to a structured approach, beginning with an in-depth exploration of the seasonal variations inherent in AHU induction motor frequencies. We then delve into the theoretical underpinnings of predictive modeling, elucidating the rationale behind our selection of Chen's Fuzzy Time Series model as the focal point of our research. Subsequently, we meticulously delineate our methodology, detailing the steps undertaken to refine and optimize the model's predictive capabilities. Our empirical analysis section provides a comprehensive overview of the experimental results, juxtaposing the performance of Chen's Fuzzy Time Series model against established seasonal prediction models. Finally, we present a cogent discussion and interpretation of the findings, offering insights into the implications and potential avenues for future research in the realm of predictive modeling for AHU induction motor frequencies.

2. Methods

The research was conducted in the library of Universitas Trilogi. This library comprises four standard zones, encompassing a combined area of 630.1 square meters. Environmental conditions in the reading room were monitored using temperature and humidity sensors. These sensors, integral to our data collection, were installed alongside a Programmable Logic Controller (PLC). A PLC is a sophisticated computer used for industrial applications, such as managing machinery or industrial processes. In this setting, the PLC was connected to the Air Handling Unit (AHU) panel, combined with an inverter. This configuration enabled precise control over the induction motor's frequency in the AHU, adjusting the air handling

system's performance based on the room's temperature and humidity levels.

The methodology of this study was structured into four distinct stages. The initial stage involved data collection and the creation of a dataset, subsequently stored in a spreadsheet file. The second stage focused on analyzing seasonal-based time divisions. The third stage entailed determining the appropriate time intervals for analysis. Finally, the last stage involved conducting experiments with modified of Chen's Fuzzy Time Series and comparing its performance against the original Chen's Fuzzy Time Series, SARIMA, Holt-Winters, SARIMAX, neural network-based ensemble model, and hybrid models of ANN and SARIMA.

2.1 Data collections

Data collection was carried out at one-minute intervals during operational hours, specifically from 07:45 to 16:45, spanning a workweek (Monday to Friday) over a period of one week. This time interval was strategically selected to provide a comprehensive understanding of occupant behaviour throughout an active semester. The collected dataset included information on the temperature, humidity, and frequency of the AHU induction motor. A scenario was established to determine the optimal temperature and humidity settings, specifically set point 22°C and 65% relative humidity. The control of the AHU induction motor based on conventional method which utilized basic reactive control. The AHU induction motor frequency starts at the minimum frequency of 0 Hz and progressively increases to a maximum value of 50 Hz. The PLC continuously receives data from the temperature and humidity sensors at intervals of 1 minute. The PLC is equipped with a logical program that monitors the temperature and humidity values. If the recorded values exceed the predetermined configuration, the PLC instructs the inverter to set the frequency to 50 Hz. Conversely, if the sensor values fall below the specified configuration, the inverter frequency is systematically reduced to 0 Hz, and vice versa. This dynamic control mechanism ensures that the AHU induction motor adjusts its frequency in response to real-time temperature and humidity conditions, optimizing the environmental parameters as required. The data collection process was facilitated by a single PLC for recording purposes. The recorded data were stored in a file format using comma-separated values.

2.2 System model

Analyzing the seasonality-based time divisions of the week in the frequency of an Air Handling Unit (AHU) induction motor dataset involves applying statistical tests to assess potential significant differences in motor frequency across different time divisions: morning, noon, and afternoon. One appropriate statistical test for this purpose is the Friedman test, a non-parametric method designed to detect differences in treatments across multiple groups [23].

1. Data preparation and null hypothesis formulation

Organize the dataset with AHU induction motor frequency measurements for each time divisions of the week. Formulate the null hypothesis (H0) that there are no significant differences in AHU induction motor frequency across time divisions. The alternative hypothesis (H1) would be that there are significant differences in AHU induction motor frequency across time divisions, indicating seasonality.

2. Rank the data and calculate Friedman statistics

Rank the AHU induction motor frequency values within each season separately, considering ties appropriately. Assign ranks based on the magnitude of AHU induction motor frequency, with lower ranks for lower frequencies. Calculate the Friedman statistic using the ranked data. The formula for the Friedman statistic is given by:

$$X^2 = \frac{12}{N(k+1)} \sum_{i=1}^{k} R_i^2 - 3(N+1) \qquad (1)$$

Where N is the total number of observations, k is the number of seasons, and R_i is the sum of ranks for the i^{-th} season.

3. Degree of freedom and critical value

To ascertain the degrees of freedom for the Friedman test, denoted as k−1, where k represents the number of seasons or time divisions, one must consider the chosen significance level, typically set at 0.05. In this specific scenario, with three distinct time divisions under examination (morning, noon, and afternoon), the calculation for degrees of freedom is 3−1, resulting in a value of 2. Therefore, for the Friedman test conducted at a significance level of 0.05 and with three time divisions, the degrees of freedom amount to 2. This parameter is crucial in the subsequent interpretation of the test results, influencing the determination of statistical

significance in the observed differences among the time divisions.

4. Hypothesis testing

Compare the calculated Friedman statistic with the critical value from the chi-square distribution. If the calculated statistic is greater than the critical value, reject the null hypothesis and conclude that there are significant differences in AHU induction motor frequency across seasons, suggesting seasonality.

2.3 Time interval analysis

This investigation delves into determining the optimal temporal resolution for monitoring Air Handling Unit (AHU) induction motor frequency, comparing four intervals: 5 minutes, 15 minutes, 30 minutes, and 1 hour. These intervals cater to different needs: the 5-minute interval for detecting immediate system variations [24], the 15-minute interval for balanced analysis [25], the 30-minute interval for capturing cyclical trends [26], and the 1-hour interval for longitudinal studies, emphasizing enduring patterns over transient fluctuations [27]. To measure the effectiveness of each time interval in capturing seasonal data, the primary method employed is entropy analysis. This analytical approach aims to discern which intervals best reflect significant changes in AHU induction motor frequency, providing the most accurate and useful data for predictive analysis [28]. Entropy, in this context, is a measure derived from information theory that quantifies the uncertainty or unpredictability of a data set. A higher entropy suggests more complexity and variability, while lower entropy indicates a more predictable and less complex dataset [29]. The formula of entropy without specific category data involves calculating the entropy of a probability distribution, where each probability represents the likelihood of an event occurring [30]. The generic formula for entropy (H) in this context is:

$$H = -\sum_{i=1}^{n} P_i \times log_2 (P_i) \quad (2)$$

Where

- H represents the entropy
- P_i is the probability of an event x occurring.
- The summation (Σ) runs over all possible events in the probability distribution.
- $log_2 (P_i)$ is the logarithm of the probability, which is typically calculated using base 2 to provide a measure in bits.
- n is the number of events or outcomes and P_i is the probability of the i^{-th} event.

2.4 Prediction model

In this study, we conducted experiments with several predictive models, each sensitive to seasonal variations. These included the Chen's Fuzzy Time Series model, its modified version, SARIMA [16], HWES-Additive [16], HWES-Multi [16], neural network (NN) based ensemble model [16], ANN-SARIMA [17], and SARIMAX [18]. The significant advancement in this research was the refinement of the defuzzification process in Chen's Fuzzy Time Series model, which was specifically engineered to integrate seasonal variations, thereby improving the precision of the predictions.

2.4.1. Chen's fuzzy time series model

Fuzzy time series is a forecasting method that combines fuzzy logic and time series analysis to handle uncertainties and imprecise information in time-related data [31]. Chen's (1996) developed a fuzzy time series based on Song and Chissom (1993) with simple operations, containing complex matrix operations, and having equal weighting to the involved elements [32]. The forecasting steps using Chen's fuzzy time series model are as follows:

1. The Universe of Discourse

Initially, the Universe of Discourse, U, is established to delineate the scope of the variables within the time series data. This delineation is pivotal for capturing the frequency range relevant to AHU induction motors.

$$U = [D_{min} - D_1; D_{max} + D_2] \quad (3)$$

Where:

- U is Universe of Discourse.
- D_{min} and D_{max} represent the smallest and largest historical data points.
- D_1 and D_2 are predetermined by the researcher to fine-tune the scope.

2. Interval Formation

In the interval formation stage, the primary goal is to establish the time intervals over which the data will be analyzed. This process is crucial for structuring the time series data into segments that can be effectively utilized in the subsequent fuzzification process. This measure is computed by taking the absolute differences between each pair of consecutive data points in the series, denoted as X_{t+1}

and X_t, and then averaging these differences across the entire series.

$$Avg = \frac{\sum_{i=1}^{n} |X_{t+1} - X_1|}{n-1} \quad (4)$$

Where:

- *Avg* is the average absolute difference.
- X_{t+1} represent the data point at time $t + 1$
- X_t represents the data point at time t.

The summation runs from $t = 1$ to $t = n - 1$, where n is the total number of data points in the series. The absolute differences capture the changes between each pair of points, and the average of these differences provides a consistent measure of the series' volatility. Once the average absolute difference is calculated, the interval length, l, is determined. This length represents the size of each interval that the universe of discourse will be divided into. Setting an appropriate interval length is crucial for capturing the underlying trends and patterns in the data without losing significant details.

$$l = \frac{Avg}{2} \quad (5)$$

The interval length is set to half of the average absolute difference. This choice strikes a balance between having too many intervals (which might overfit the noise in the data) and too few (which might miss important variations). Finally, the number of intervals, or fuzzy numbers, denoted as p, is calculated. This number determines how many distinct fuzzy sets will be used to represent the data.

$$p = \frac{(D_{max} + D_2 - D_{min} - D_1)}{l} \quad (6)$$

The numerator of this formula, $D_{max} + D_2 - D_{min} - D_1$, represents the adjusted range of the data, and dividing this by the interval length, l, yields the total number of intervals.

3. Fuzzification

Fuzzification is a crucial step in processing time series data within Chen's Fuzzy Time Series model. This process involves converting precise numerical data into fuzzy sets to effectively manage the inherent uncertainty present in real-world data. Each data point in the series is associated with a membership function, which reflects its degree of belonging to various fuzzy sets. These fuzzy sets are classes or groups with a continuum of membership degrees, allowing for a more nuanced representation of data. Consider U as the universal set encompassing all possible values in the time series, represented as $U = \{u_1, u_2, \ldots, u_n\}$ where each u_i is a distinct value in U. The linguistic variable A_i regarding U is formulated using equation below.

$$A_i = \frac{\mu_{Ai}(u_1)}{u_1} + \frac{\mu_{Ai}(u_{21})}{u_{12}} + \frac{\mu_{Ai}(u_{n1})}{u_n} \quad (7)$$

In this equation, μ_{Ai} denotes the membership function of the fuzzy set A_i, mapping each element in U to a membership value within the interval [0, 1]. The function quantifies the degree to which each u_j belongs to the fuzzy set A_i. The assignment of membership degrees to each u_j in A_i is governed by specific rules based on the actual data point X_t.

a. If X_t is equal to u_i: The degree of membership for u_i is assigned a value of 1, signifying full membership. The immediate subsequent value u_{i+1} are given a membership degree of 0.5, reflecting a partial membership. All other values are assigned a membership degree of 0, indicating no membership.
b. If X_t falls within the range of ($1 \leq i \leq p$): The membership degree for u_i is set to 1. The adjacent values, u_{i-1} and u_{i+1}, receive a membership degree of 0.5, denoting partial membership. Values outside this range are declared to have zero membership.
c. If X_t is specifically included in u_i: The degree of membership for u_i is 1, and for u_{i-1}, the membership is considered zero.

4. Establish Fuzzy Relationship

After the data are transformed into fuzzy sets via fuzzification, the next task is to determine the Fuzzy Logic Relations (FLR) between these sets. An FLR essentially captures a relationship between two consecutive data points (or periods) in the fuzzy domain. It's expressed in the form $A_i \rightarrow A_j$, indicating a transition or a link from one fuzzy set A_i to another fuzzy set A_j. To determine an FLR, one looks at the fuzzified values obtained from the previous step. If, for instance, the fuzzification process for month n yields a particular fuzzy set Ai, and for month $n + 1$ it yields Aj, then the FLR is noted as $A_i \rightarrow A_j$. This process is repeated for the entire time series, creating a chain of FLRs that reflect the dynamic nature of the data. Once individual FLRs are identified, they are grouped into Fuzzy Logic Relations Groups (FRLG). An FRLG is a collection of FLRs that share the same starting point. For example, if over several months, the fuzzy relation A_1 transitions to A_2, back to A_1, then to A_3, and again to A_1, the FLRG for the starting point A_1 would be

$A_1 \rightarrow A_{1,2}, A_3$. This grouping helps in understanding the various paths that the system can take from a particular state, reflecting the inherent uncertainties and providing a comprehensive view of potential future trends.

5. Defuzzification

Defuzzification in Chen's fuzzy time series model involves converting fuzzy sets into a single crisp numerical value, which can be used for making precise and actionable forecasts. This process is governed by several rules depending on the relationships between the fuzzy sets identified during the fuzzification stage. Several rules are established in the defuzzification phase of Chen's fuzzy time series model to ensure the translation of fuzzy sets into numerical predictions. Initially, when a fuzzy set A_i is yielded from the fuzzification at time period t without any subsequent fuzzy logical relationships, denoted as $A_i \rightarrow \emptyset$, it is considered isolated. The forecasted value for the next period F_{t+1} is determined by the middle value, m_i, of the interval u_i, where the maximum of A_i membership function resides. In instances where a single fuzzy logical relationship (FLR) exists between the fuzzified set A_i and another set A_j expressed as $A_i \rightarrow A_j$ attention is shifted to the interval u_j. The middle value m_j representing the maximum of A_j membership function is adopted as the forecast for F_{t+1}. A more complex rule is applied when multiple FLRs are identified, linking A_i to several sets, shown as $A_i \rightarrow A_{j1}, A_{j2}, \ldots,$. The forecast for F_{t+1} in such cases is calculated as the average of the middle values of the intervals $u_{j1}, u_{j2}, \ldots, u_{jk}$, represented by $m_{j1}, m_{j2}, \ldots, m_{jk}$. The formula applied is:

$$F_{t+1} = \frac{m_{j1} + m_{j2} + \cdots + m_{jk}}{k} \tag{8}$$

The value of k is identified as the count of middle values(midpoints). This count reflects the number of sets involved when multiple Fuzzy Logical Relationships (FLRs) are present. To ascertain the middle value m_i of any interval, the following formula is utilized:

$$m_i = \frac{(T_i + B_i)}{2} \tag{9}$$

T_i represents the top value of the interval, and B_i signifies the bottom value. By averaging these two, the precise middle value of the interval is calculated, serving as a critical component in the defuzzification process of Chen's fuzzy time series model.

2.4.2. Proposed method

In this research, we have developed an enhanced version of Chen's fuzzy time series model by introducing an exponential defuzzification method. This method integrates two parameters, alpha (α) and beta (β), into the standard defuzzification rules. The underlying concept is inspired by the Exponential Moving Average (EMA) approach, which employs α and β to account for seasonal variations [33]. The incorporation of α and β is designed to bring a higher degree of adaptability and responsiveness to the defuzzification process. Specifically, α is used to modulate the influence of recent data, enhancing the model's adaptability, while β is applied to adjust the significance of historical data, thereby ensuring the model remains responsive. This innovative approach finely tunes the defuzzification process, effectively balancing recent and historical data to more accurately reflect trends and seasonal changes in the data. The exponential defuzzification is mathematically expressed as:

$$EF_t = (\alpha \times C_n) + (\beta \times F_t) \tag{10}$$

Where:

- EF_t signifies the value of exponential defuzzification
- F_t is the forecast value obtained from the initial defuzzification

This formula, leveraging alpha (α) and beta (β), is meticulously crafted to evenly distribute the influence of recent and historical data, providing a robust and temporally sensitive enhancement to the standard model.

2.5 Performance measurement

Performance measurement utilizes Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R-Square. MAPE is a widely used metric to assess forecasting model accuracy, measuring the average absolute percentage difference between predicted and actual values [34]. A lower MAPE suggests better accuracy, with 0% indicating a perfect forecast. However, MAPE may have limitations, particularly with small or zero actual values. RMSE computes the square root of the average squared differences between predicted and actual values [35]. Similar to MAPE, a lower RMSE signifies better accuracy. RMSE places more weight on larger errors compared to MAPE, making it sensitive to outliers and providing an overall measure of model

prediction error. R-Square, a coefficient of determination from regression analysis, maximizes the best coefficient value to include in the model [36]. It indicates how well the model explains outcome variability and should be used alongside other metrics and domain knowledge for evaluation.

3. Result and discussion

This study utilized a dataset collected from a reading room in the library at Universitas Trilogi, focusing on three main variables: temperature, relative humidity, and the frequency of the AHU induction motor. Data collection occurred during an active week in the semester, with observations recorded every minute, resulting in 480 daily records and a total of 2400 over the week. The one-minute interval was selected to capture the dynamic relationship among temperature, humidity, and the AHU's induction motor frequency accurately. Data collection employed a conventional method using basic reactive controls operated by a PLC to regulate the inverter and determine the AHU induction motor frequency. The dataset was divided into four subsets corresponding to 5 minutes, 15 minutes, 30 minutes, and 1-hour intervals.

3.1 Seasonality-based time divisions

The application of the Friedman test revealed significant findings, with a computed test statistic (Q) of 17.033. This value, representing the sum of ranks for differences among time intervals, determines whether these intervals significantly influence AHU induction motor frequency data. The test resulted in a remarkably low p-value of 0.0002001, indicating the probability of obtaining test results as extreme as those observed, assuming the null hypothesis is true. With a predefined significance level (α) of 0.05, the obtained p-value (0.0002001 < 0.05) led to the confident rejection of the null hypothesis. This rejection indicates significant differences in the effects of various time intervals on frequency data. These findings, summarized in Table 1, provide statistical evidence supporting the conclusion regarding the selection of time intervals.

Table 1. Summary of Friedman Test Results

Parameter	Value
Test Statistic (Q)	17.033
Degrees of Freedom (df)	2
p-value	0.0002001
Significance Level (α)	0.05
Significant Difference	Yes

Table 2. Time Interval Results

Time Interval	Entropy
5 min	130.638
15 min	32.677
30 min	12.749
1 h	5.965

The rejection of the null hypothesis indicates a significant difference in motor frequency among the morning, noon, and afternoon time divisions. The calculated p-value, much smaller than the predetermined significance level, strengthens the conclusion that observed frequency variations are unlikely to be random. Thus, the data provides statistical evidence supporting the assertion of a significant distinction in motor frequency across the specified time divisions. The Friedman test applied to the seasonality-based time division analysis highlights a statistically significant difference in the frequency of the Air Handling Unit (AHU) induction motor during morning, noon, and afternoon periods.

3.2 Proper time interval

The current study aimed to determine the optimal time interval for analyzing datasets, specifically by utilizing the concept of entropy. Entropy serves as a metric for quantifying the level of disorder or uncertainty within the dataset. Various time intervals—5 minutes, 15 minutes, 30 minutes, and 1 hour—were explored to cover the entire dataset for distinct periods within a week. Detailed attention was given to entropy analysis at each of these time intervals. The results of this analysis, crucial for identifying the optimal interval for data analysis, are documented in Table 2.

The entropy value of 5.965 indicated the level of unpredictability within the environmental conditions. Notably, the study suggested that the optimal time period for this analysis was 1 hour, as indicated by the entropy value exceeding the minimum criterion of 5.965, signifying a robust goodness of fit. This longer timeframe allows for a comprehensive understanding of the AHU induction motor frequency dynamics under the defined ambient parameters.

3.3 Prediction of AHU induction motor frequency

The prediction of AHU induction motor frequency involved various models, encompassing proposed model with state-of-the-art methods

includes SARIMA [16], HWES-Additive [16], HWES-Multi [16], neural network (NN) based ensemble model [16], ANN-SARIMA [17], and SARIMAX [18]. The distinct model fit parameters for each prediction approach are detailed in Tables 3 through 6. The parameters for the neural network (NN) based ensemble model and the ANN-SARIMA hybrid model across all time intervals to ensure consistency and facilitate meaningful comparisons. Parameters included the Rectified Linear Unit (ReLU) activation function, 10 epochs, a batch size of 32, the Adam optimizer, and binary crossentropy loss. ReLU was chosen for its ability to capture complex patterns, while 10 epochs struck a balance between convergence and computational efficiency. A batch size of 32 ensured stable and efficient training and the Adam optimizer's adaptive learning rate accelerated convergence. The binary crossentropy loss function was suited for binary classification tasks.

Table 3 presents the model fit parameters for the Modified Chen method, with Alpha (α) and Beta (β) controlling trend and seasonality smoothing, respectively. These parameters remain constant across all intervals, ensuring consistent data smoothing. In Table 4, SARIMA, ANN-SARIMA, and SARIMAX models are detailed with parameters such as p, d, q, P, D, Q, and s, allowing adaptation to varying data patterns across intervals. Tables 5 and 6 outline HWES-Additive and HWES-Multiplicative parameters, including Alpha (α) and Gamma (γ) for level and trend smoothing. Their variability across intervals underscores the models' adaptability to diverse data trends.

Table 3. Model Fit Parameter of Modified Chen

Parameter	5 min	15 min	30 min	1 h
Alpha(α)	0.1	0.1	0.1	0.9
Beta(β)	0.9	0.9	0.9	0.1

Table 4. Model Fit Parameter of SARIMA, ANN-SARIMA, and SARIMAX

Parameter	5 min	15 min	30 min	1 h
p,d,q	2,0,0	2,0,0	0,0,0	0,0,1
P,D,Q,s	1,0,0,108	1,0,0,36	1,0,1,18	1,0,2,9

Table 5. Model Fit Parameter of HWES-Additive

Parameter	5 min	15 min	30 min	1 h
Alpha(α)	0.0552	0.1458	1.4901	1.7109
Gamma(γ)	6.3753	0.0001	2.2606	0

Table 6. Model Fit Parameter of HWES-Multiplicative

Parameter	5 min	15 min	30 min	1 h
Alpha(α)	0.0527	0.1317	1.4901	1.4901
Gamma(γ)	0.0007	6.7937	0	7.2505

The performance metrics, including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R-squared values, served as quantifiable measures of each model's accuracy and predictive capability. Building upon the insights gained from entropy analysis, the study evaluated the forecasting models' performance using MAPE across the same time intervals, as shown in Figure 1. MAPE quantifies the average absolute percentage deviation between predicted and actual values, serving as a key indicator of a model's accuracy and reliability.

The results suggest that the Modified Chen's model consistently outperforms other methods, with low MAPE values ranging from 2.405 to 4.699 across all time intervals, indicating superior accuracy in predicting AHU induction motor frequency. Traditional SARIMA models show comparatively higher MAPE values (23.462 to 36.197), indicating less accuracy. The Holt-Winters Exponential Smoothing (HWES) models, both additive and multiplicative, perform relatively better than SARIMA but are surpassed by the Modified Chen's model. The neural network (NN) ensemble model performs well for shorter intervals (5-minute and 30-minute), with MAPE values of 19.824 and 18.468, respectively, but struggles for longer intervals (1-hour) with a high MAPE of 79.622. The ANN-SARIMA hybrid model demonstrates exceptional accuracy (MAPE: 0.294 to 0.459), highlighting the efficacy of combining neural networks with SARIMA. The SARIMAX model

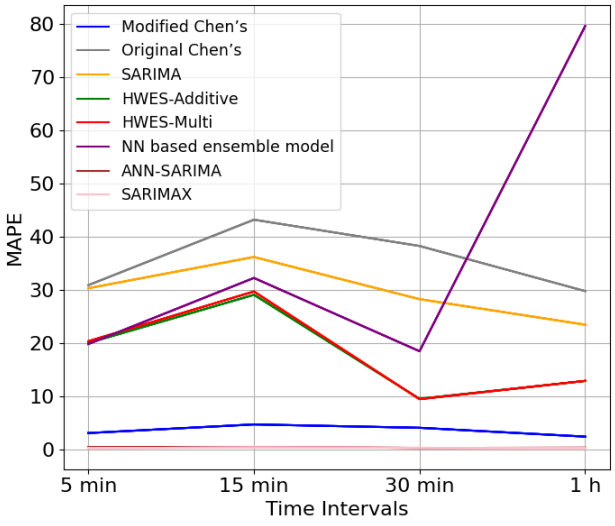

Figure. 1 Score of MAPE

also exhibits impressive accuracy (MAPE: 0.256 to 0.362) across all intervals.

The SARIMAX model, an extension of SARIMA, exhibits remarkable predictive capabilities in forecasting AHU induction motor frequency. Analysis of MAPE values indicates that Modified Chen's model, ANN-SARIMA, and SARIMAX are the most accurate forecasting methods across different time intervals. Conversely, traditional SARIMA approaches tend to yield higher prediction errors compared to these models. While the NN-based ensemble model shows promise, further refinement may be necessary for long-term forecasting.

The RMSE values, presented in Figure 2, provide another layer of understanding by quantifying the square root of the average squared differences between the predicted and actual values. This metric is particularly sensitive to large errors, making it a robust measure of predictive accuracy.

The RMSE analysis reveals distinct patterns among the forecasting models. The Modified Chen's Model consistently demonstrates superior accuracy, with RMSE values ranging from 0.724 to 1.474 across all time intervals. In contrast, traditional methods like SARIMA and the Original Chen's Model exhibit higher RMSE values, ranging from 7.806 to 14.1163, indicating less precision in predicting AHU induction motor frequency.

HWES models, both Additive and Multiplicative, perform competitively, displaying relatively low RMSE values across various time intervals. However, the NN-based Ensemble Model shows higher RMSE values, especially for longer forecasting horizons, highlighting potential limitations in its predictive capabilities. Similarly,

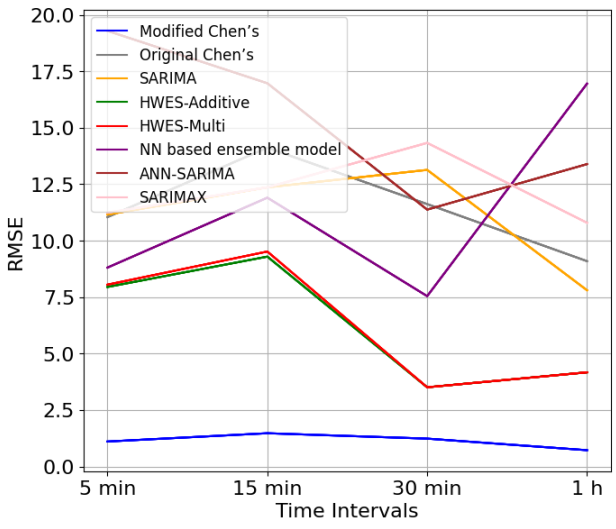

Figure. 2 Score of RMSE

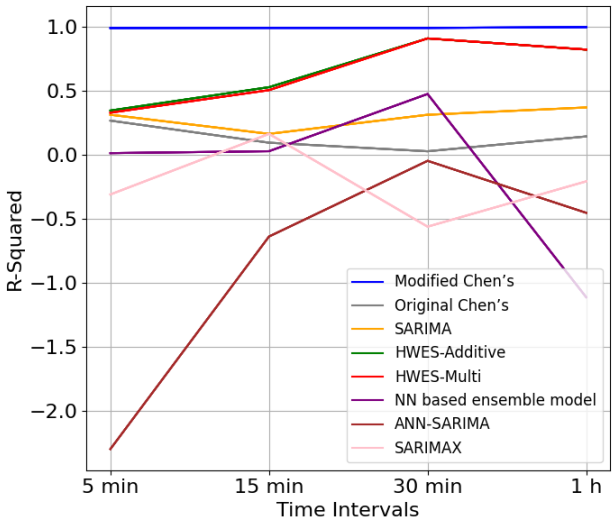

Figure. 3 Score of R-Squared

the ANN-SARIMA Hybrid Model demonstrates mixed performance, with higher RMSE values for extended forecasting periods.

Comparatively, the SARIMAX Model yields RMSE values akin to SARIMA, albeit slightly higher for longer forecasting intervals. Overall, lower RMSE values signify enhanced accuracy and reliability in predicting AHU induction motor frequency. The Modified Chen's Model emerges as the most precise predictor, while traditional and innovative methods exhibit varying predictive performance across time intervals.

The R-squared values, as shown in Table 9, depict the proportion of the variance in the dependent variable that is predictable from the independent variables. R-squared values range from 0 to 1, with higher values indicating a model that can better account for the variability in the dataset and thus a better fit.

The R-squared values for Chen's model are The R-squared values presented in Table 9 provide insights into the predictive performance and explanatory power of the various forecasting models for AHU induction motor frequency. R-squared, also known as the coefficient of determination, is a measure of the proportion of variance in the dependent variable (motor frequency) that is explained by the independent variables (predictors) included in the model. Across the different time intervals, the Modified Chen's Model consistently exhibits high R-squared values, ranging from 0.987 to 0.994. These values indicate that the Modified Chen's Model can account for approximately 98.7% to 99.4% of the variance in AHU induction motor frequency, suggesting a strong and reliable fit to the data.

In contrast, traditional approaches like SARIMA and the Original Chen's Model yield comparatively lower R-squared values, ranging from 0.0264 to 0.368. These models are less effective at explaining the variability in motor frequency data, indicating weaker fits. The Holt-Winters Exponential Smoothing (HWES) models, both Additive and Multiplicative, demonstrate moderate to high R-squared values across different time intervals, ranging from 0.327 to 0.906. These values suggest that HWES models provide reasonable explanations for the observed variability in AHU induction motor frequency. In contrast, the NN-based Ensemble Model and the ANN-SARIMA Hybrid Model display mixed performance, with R-squared values varying widely across time intervals. While some intervals show positive R-squared values, indicating a reasonable fit, others exhibit negative values, suggesting poor model performance or overfitting. The SARIMAX Model also displays mixed results, with some intervals showing negative R-squared values, indicating that the model fails to explain the variability in motor frequency data effectively.

The present study underscores the supremacy of the modified Chen's model in predicting AHU induction motor frequency through the use of advanced predictive modeling techniques. The near-perfect R-squared values achieved by the modified Chen's model reflect its exceptional ability to explain the variance within the dependent variable, surpassing many contemporary models. Although the SARIMA, HWES, ANN-SARIMA, SARIMAX, and neural network (NN) ensemble models are well-regarded in the field, their performance pales in comparison to the nuanced understanding achieved by the modified Chen's model. The modified Chen's model stands out by incorporating sophisticated algorithmic adjustments that enable it to delve deeper into the dataset and identify subtle patterns and fluctuations that simpler models might overlook. This attribute is particularly crucial in understanding and predicting the dynamics of AHU induction motor frequency, where data intricacies abound. While the ANN-SARIMA hybrid model and SARIMAX exhibit promising results, the neural network (NN) ensemble model shows mixed performance, with significant challenges noted for longer forecasting horizons.

The practical applicability of the modified Chen's model is underscored by its adaptability and consistency across various time intervals. This attribute, combined with its ability to maintain accuracy and reliability irrespective of the granularity of the data, makes it a valuable tool in dynamic conditions prevalent in practical settings. Moreover, its precision and accuracy are not compromised, even when the data collection frequency and demand are diverse, highlighting its effectiveness in real-world scenarios.

Furthermore, the study's findings that a 1-hour time interval is most effective, as indicated by the low entropy value of 5.965, highlight the model's capacity to capture comprehensive dynamics. This proficiency is crucial in enhancing the prediction accuracy for the AHU induction motor frequency. The model's robust alignment with actual data, deep insight into complex dynamics, and superior performance across multiple statistical metrics notably elevate its precision above other models. The consistent application of alpha (α) and beta (β) values in the model contributes significantly to its stability and effectiveness. These parameters allow the model to adapt to trends over varying time frames while minimizing prediction errors. They play a crucial role in avoiding issues like overfitting or underfitting, which can compromise the model's effectiveness.

4. Conclusion

The study aimed to identify the optimal time interval for analyzing AHU induction motor frequency datasets by employing entropy analysis and evaluating the performance of various forecasting models. Entropy served as a metric to quantify the level of disorder or uncertainty within the dataset, with results indicating a decreasing trend in entropy with longer time intervals. The optimal time interval of 1 hour was identified based on the entropy analysis, suggesting a comprehensive understanding of AHU induction motor frequency dynamics within defined ambient parameters. The evaluation of forecasting models using Mean Absolute Percentage Error (MAPE) revealed that the modified Chen's model consistently outperformed other methods across all time intervals, demonstrating superior accuracy in predicting AHU induction motor frequency. The MAPE values for the modified Chen's model ranged from 2.405 to 4.699, indicating minimal absolute percentage deviation between predicted and actual values. Conversely, traditional SARIMA models exhibited higher MAPE values (ranging from 23.462 to 36.197), reflecting less accuracy in prediction. The Holt-Winters Exponential Smoothing (HWES) models showed relatively better performance than SARIMA but were surpassed by the modified Chen's model. Notably, the neural network (NN) ensemble model displayed promising results for shorter intervals but struggled with longer intervals,

indicating potential challenges in long-term forecasting. The ANN-SARIMA hybrid model demonstrated exceptional accuracy across all intervals, underscoring the effectiveness of combining neural networks with SARIMA. Similarly, the SARIMAX model exhibited impressive accuracy across all intervals, reflecting its remarkable predictive capabilities.

Furthermore, the Root Mean Squared Error (RMSE) analysis provided insights into the precision of forecasting models. The modified Chen's model consistently demonstrated superior accuracy, with lower RMSE values ranging from 0.724 to 1.474 across all time intervals. In contrast, traditional methods like SARIMA and the original Chen's model exhibited higher RMSE values, indicating less precision in predicting AHU induction motor frequency. The HWES models, both Additive and Multiplicative, performed competitively, while the NN-based Ensemble Model and the ANN-SARIMA hybrid model showed mixed performance across different forecasting horizons. Moreover, the R-squared values elucidated the predictive performance and explanatory power of forecasting models. The modified Chen's model consistently exhibited high R-squared values, indicating a strong and reliable fit to the data, while traditional approaches like SARIMA showed comparatively lower values.

The quantitative outcomes substantiate the superior performance of the modified Chen's model, especially its capacity to deliver precise predictions with limited data. The comparative data unequivocally establish the model's superiority in terms of accuracy and reliability when contrasted with baseline and state-of-the-art methodologies. Through exhaustive testing and comparative evaluations, the modified Chen's model has demonstrated its prowess in forecasting accuracy, thus setting a new standard for future investigations and practical implementations in environmental control systems. Subsequent research endeavors will concentrate on integrating advanced machine learning techniques into the modified Chen model to elevate its predictive capabilities further. Additionally, conducting a more extensive comparative assessment with other sophisticated models not covered in this study will enrich our comprehension of relative performances. The model's application across diverse datasets will serve to evaluate its robustness and adaptability across various HVAC systems and operational environments.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this research article. No personal circumstances or interests exist that could be perceived as inappropriately influencing the representation or interpretation of the reported research results.

Author Contributions

Conceptualization, Yaddarabullah and Yoga Alviando; methodology, Dewi Lestari; software, Yoga Alviando; validation, Yaddarabullah, Aedah Binti Abd Rahman, and Amna Saad; formal analysis, Dewi Lestari; investigation, Yaddarabullah; resources, Yoga Alviando; data curation, Yoga Alviando; writing—original draft preparation, Yaddarabullah; writing—review and editing, Yaddarabullah; visualization, Yaddarabullah; supervision, Aedah Binti Abd Rahman, and Amna Saad; project administration, Yaddarabullah; funding acquisition, Yaddarabullah.

Acknowledgments

We would like to express our sincere gratitude to all those who contributed to the successful completion of this research. First and foremost, we extend our heartfelt appreciation to our supervisor at Asia e University (AeU) and Universiti Kuala Lumpur (UniKL), Prof. Dr. Aedah Binti Abd Rahman and Ts. Dr. Amna Saad, for their invaluable guidance, encouragement, and insightful feedback throughout the project. We would also like to thank Indonesian Ministry of Education, Culture, Research and Technology for giving research funding and Universitas Trilogi for giving permission to use the library as the object of research.

References

[1] Yang, G. Hu, and C. J. Spanos, “HVAC Energy Cost Optimization for a Multizone Building via a Decentralized Approach”, *IEEE Trans. Autom. Sci. Eng.*, Vol. 17, No. 4, pp. 1950–1960, 2020, doi: 10.1109/TASE.2020.2983486.

[2] W. Jung and F. Jazizadeh, “Human-in-the-loop HVAC operations: A quantitative review on occupancy, comfort, and energy-efficiency dimensions”, *Appl. Energy*, Vol. 239, pp. 1471–1508, 2019, doi: https://doi.org/10.1016/j.apenergy.2019.01.070.

[3] W. Li, C. Koo, T. Hong, J. Oh, S. H. Cha, and S. Wang, “A novel operation approach for the energy efficiency improvement of the HVAC system in office spaces through real-time big

data analytics", *Renew. Sustain. Energy Rev.*, Vol. 127, p. 109885, 2020, doi: https://doi.org/10.1016/j.rser.2020.109885.

[4] G. Balasubramanian, P. Chandrasekar, and S. A. Alexandar, "Variable Frequency Drive operated Air Blower in Air Handling Unit of Heating, Ventilation and Air Conditioning Systems", In: *Proc. of 2022 IEEE Delhi Section Conference (DELCON)*, pp. 1–7, 2022, doi: 10.1109/DELCON54057.2022.9753386.

[5] A. Kaushik, M. Arif, P. Tumula, and O. J. Ebohon, "Effect of thermal comfort on occupant productivity in office buildings: Response surface analysis", *Build. Environ.*, Vol. 180, p. 107021, 2020, doi: https://doi.org/10.1016/j.buildenv.2020.107021.

[6] T. Park, B. Kim, G. Hwang, Y. Kang, I. Lee, and Y. Ahn, "Improving Comfort and Air Conditioner Performance by Optimizing Controllers under Actual Usage Conditions", *Applied Sciences*, Vol. 11, No. 11. 2021. doi: 10.3390/app11114818.

[7] X. Lu, V. Adetola, and Z. O'Neill, "What are the impacts on the HVAC system when it provides frequency regulation? – A comprehensive case study with a Multi-Zone variable air volume (VAV) system", *Energy Build.*, Vol. 243, p. 110995, 2021, doi: 10.1016/j.enbuild.2021.110995.

[8] S. Walgama, S. Kumarawadu, and C. D. Pathirana, "Performance Analysis of Modified Elliot Function Based Temperature Controller in the Heating Mode Operation of AHU", In: *Proc. of 2021 Moratuwa Engineering Research Conference (MERCon)*, pp. 89–94, 2021, doi: 10.1109/MERCon52712.2021.9525770.

[9] N. M.-A. Mutombo and B. P. Numbi, "Development of a Linear Regression Model Based on the Most Influential Predictors for a Research Office Cooling Load", *Energies*, Vol. 15, No. 14, 2022, doi: 10.3390/en15145097.

[10] Zhang, Dongxu, Huang, Kailiang, Feng, Guohui, Lv, Xue, and Xu, Jintian, "Study on fresh air load characteristics and energy saving measures of low energy consumption buildings in severe cold area", *E3S Web Conf.*, Vol. 356, p. 1059, 2022, doi: 10.1051/e3sconf/202235601059.

[11] S. A. Sulaiman and A. H. Hassan, "Analysis of Annual Cooling Energy Requirements for Glazed Academic Buildings", *J. Appl. Sci.*, Vol. 11, pp. 2024–2029, 2011, doi: https://doi.org/10.3923/JAS.2011.2024.2029.

[12] Q. M. Ali, M. J. Thaheem, F. Ullah, and S. M. E. Sepasgozar, "The Performance Gap in Energy-Efficient Office Buildings: How the Occupants Can Help?", *Energies*, 2020, doi: 10.3390/en13061480

[13] J. Nuñez Alvarez, R. Linares Vicente, Y. Zamora, E. Noriega Angarita, F. Rodríguez, and L. Mendoza, "Automation of the air conditioning system for aseptic rooms of pharmaceutical production", *Int. J. Power Electron. Drive Syst.*, Vol. 14, pp. 1479–1488, 2023, doi: 10.11591/ijpeds.v14.i3.pp1479-1488.

[14] S. Nadweh, Z. Barakat, and G. Hayek, "EMC Installation for Variable Speed Drive Systems (VSDs): Fields, Emissions, Coupling, and Shielding", *Handbook of Research on Smart Power System Operation and Control*, pp. 308–329, 2019, doi: 10.4018/978-1-5225-8030-0.ch013.

[15] S. Wilson, S. Maalej, and A. Hunter, "Increasing the Efficiency of HVAC Systems using Schedule-Based Control", In: *Proc. of 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 238–243, 2022, doi: 10.1109/UEMCON54665.2022.9965717.

[16] J. G. Jetcheva, M. Majidpour, and W.-P. Chen, "Neural network model ensembles for building-level electricity load forecasts", *Energy Build.*, Vol. 84, pp. 214–223, 2014, doi: https://doi.org/10.1016/j.enbuild.2014.08.004.

[17] L. Xuemei, D. Lixing, S. Ming, X. Gang, and L. Jibin, "A Novel Air-Conditioning Load Prediction Based on ARIMA and BPNN Model", In: *Proc. of 2009 Asia-Pacific Conference on Information Processing*, pp. 51–54, 2009, doi: 10.1109/APCIP.2009.21.

[18] M. Fathollahzadeh, "Integrated framework for modeling and optimization of commercial district cooling systems", *Colorado School of Mines. Arthur Lakes Library*, 2021. [Online]. Available: https://hdl.handle.net/11124/176506

[19] B. Yildiz, J. Bilbao, and A. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting", *Renew. Sustain. Energy Rev.*, Vol. 73, pp. 1104–1122, 2017, doi: 10.1016/j.rser.2017.02.023.

[20] M. Khashei, M. Bijari, and S. R. Hejazi, "Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting", *Soft Comput.*, Vol. 16, pp. 1091–1105, 2012, doi: 10.1007/s00500-012-0805-9.

[21] Q. Yuan, J. Shang, X. Cao, C. Zhang, X. Geng, and J. Han, "Detecting Multiple Periods and Periodic Patterns in Event Time Sequences,"

In: *Proc. of 2017 ACM Conf. Inf. Knowl. Manag.*, 2017, doi: 10.1145/3132847.3133027.

[22] B. R. Carter *et al.*, "Time of Detection as a Metric for Prioritizing Between Climate Observation Quality, Frequency, and Duration", *Geophys. Res. Lett.*, Vol. 46, pp. 3853–3861, 2019, doi: https://doi.org/10.1029/2018GL080773

[23] J. C. W. Rayner and G. Livingston Jr, "Relating the Friedman test adjusted for ties, the Cochran–Mantel–Haenszel mean score test and the ANOVA F test", *Commun. Stat. - Theory Methods*, Vol. 52, No. 12, pp. 4369–4378, 2023, doi: 10.1080/03610926.2021.1994606.

[24] A. K. Pandey and P. Pandey, "Systematic approach to proactive detection & analysis of response time variations in multi-tier systems", In: *Proc. of 2016 IEEE Annu. India Conf.*, pp. 1–5, 2016, [Online]. Available: https://api.semanticscholar.org/CorpusID:2845390

[25] G. Christodoulou, "Efficient and Scalable Management of Interval Data," in *EDBT/ICDT Workshops*, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:258559077

[26] R. May, C. Ml, A. J. Richardson, S. Kinsella, A. Khalil, and D. N. Lucas, "Defining the Decision-to-Delivery Interval at Cesarean Section: Narrative Literature Review and Proposal for Standardization", *Obstetric Anesthesia Digest*. 2022. doi: 10.1097/01.aoa.0000891744.46670.3e.

[27] J. V Jasien *et al.*, "Comparison of Extraocular and Intraocular Pressure Transducers for Measurement of Transient Intraocular Pressure Fluctuations Using Continuous Wireless Telemetry", *Scientific Reports*. 2020. doi: 10.1038/s41598-020-77880-8.

[28] M. Ausloos, O. Nedić, and A. Dekanski, "Seasonal Entropy, Diversity and Inequality Measures of Submitted and Accepted Papers Distributions in Peer-Reviewed Journals", *Entropy*, Vol. 21, 2019, doi: https://doi.org/10.3390/e21060564

[29] F. Xiong, S. Guo, L. Chen, F.-J. Chang, Y. Zhong, and P. Liu, "Identification of flood seasonality using an entropy-based method", *Stoch. Environ. Res. Risk Assess.*, Vol. 32, pp. 3021–3035, 2018, doi: 10.1007/s00477-018-1614-1

[30] A. J. López-Menéndez and R. Pérez-Suárez, "Entropy Application for Forecasting", *Entropy*, Vol. 22, No. 6, 2020, doi: 10.3390/e22060604.

[31] P. C. de L. Silva, H. J. Sadaei, R. Ballini, and F. G. Guimarães, "Probabilistic Forecasting With Fuzzy Time Series", *IEEE Trans. Fuzzy Syst.*, Vol. 28, No. 8, pp. 1771–1784, 2020, doi: 10.1109/TFUZZ.2019.2922152.

[32] Zaenurrohman, S. Hariyanto, and T. Udjiani, "Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting", *J. Phys. Conf. Ser.*, Vol. 1943, No. 1, p. 12128, 2021, doi: 10.1088/1742-6596/1943/1/012128.

[33] D. Haynes, S. M. Corns, and G. K. Venayagamoorthy, "An Exponential Moving Average algorithm", In: *Proc. of 2012 IEEE Congr. Evol. Comput.*, pp. 1–8, 2012, doi: https://doi.org/10.1109/CEC.2012.6252962

[34] Al-Khowarizmi, S. Efendi, M. K. Nasution, and M. Herman, "The Role of Detection Rate in MAPE to Improve Measurement Accuracy for Predicting FinTech Data in Various Regressions", In: *Proc. of 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, pp. 874–879, 2023, doi: 10.1109/ICCoSITE57641.2023.10127820.

[35] S. Prayudani, A. Hizriadi, Y. Lase, Y. Fatmi, and A. Al-Khowarizmi, "Analysis Accuracy Of Forecasting Measurement Technique On Random K-Nearest Neighbor (RKNN) Using MAPE And MSE", *J. Phys. Conf. Ser.*, Vol. 1361, p. 12089, 2019, doi: 10.1088/1742-6596/1361/1/012089.

[36] A. R. Sánchez, R. S. Gómez, and C. B. G. García, "The coefficient of determination in the ridge regression", *Commun. Stat. - Simul. Comput.*, Vol. 51, pp. 201–219, 2019, doi: 10.1080/03610918.2019.1649421.