

A LSTM algorithm-driven deep learning approach to estimating repair and maintenance costs of apartment buildings

Engineering,
Construction and
Architectural
Management

369

Ji-Myong Kim

*Department of Architectural Engineering, Mokpo National University,
Muan, South Korea*

Sang-Guk Yum and Manik Das Adhikari

*Department of Civil Engineering, Gangneung-Wonju National University,
Gangneung, South Korea, and*

Junseo Bae

Division of Smart Cities, Korea University – Sejong Campus, Sejong, South Korea

Received 27 November 2023

Revised 20 May 2024

30 June 2024

Accepted 4 August 2024

Abstract

Purpose – This study proposes a deep learning algorithm-based model to predict the repair and maintenance costs of apartment buildings, by collecting repair and maintenance cost data that were incurred in an actual apartment complex. More specifically, a long short-term memory (LSTM) algorithm was adopted to develop the prediction model, while the robustness of the model was verified by recurrent neural networks (RNN) and gated recurrent units (GRU) models.

Design/methodology/approach – Repair and maintenance cost data incurred in actual apartment complexes is collected, along with various input variables, such as repair and maintenance timing (calendar year), usage types, building ages, temperature, precipitation, wind speed, humidity and solar radiation. Then, the LSTM algorithm is employed to predict the costs, while two other learning models (RNN and GRU) are taught to validate the robustness of the LSTM model based on *R*-squared values, mean absolute errors and root mean square errors.

Findings – The LSTM model's learning is more accurate and reliable to predict repair and maintenance costs of apartment complex, compared to the RNN and GRU models' learning performance. The proposed model provides a valuable tool that can contribute to mitigating financial management risks and reducing losses in forthcoming apartment construction projects.

Originality/value – Gathering a real-world high-quality data set of apartment's repair and maintenance costs, this study provides a highly reliable prediction model that can respond to various scenarios to help apartment complex managers plan resources more efficiently, and manage the budget required for repair and maintenance more effectively.

Keywords Construction, Project management, Methodology, Decision support systems, Approach

Paper type Research paper

1. Introduction

It is acknowledged that it is challenging to build individual houses specifically in densely populated large urban areas. To overcome this hurdle, apartment complexes have become

© Ji-Myong Kim, Sang-Guk Yum, Manik Das Adhikari and Junseo Bae. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

This research was funded by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2022R1F1A106314112). Also, this work was supported by a Korea University Grant.



one of the most-widely used residential facility types that offer a practical solution to accommodate tons of people in such limited urban areas, while providing more efficient housing options (Islam *et al.*, 2021). Given the complexity and scale of apartments, a sophisticated and scientific facilities management is crucial to meet with many residents' satisfaction. Facilities management (FM) plays a pivotal role in ensuring the efficient operation, maintenance and improvement of physical assets within an organization or property (Bröchner *et al.*, 2019). Effective facilities management, for example based on the most-feasible budgets, helps facility managers reduce operating costs through proper maintenance, energy efficiency and the optimal utilization of resources. To this end, regular maintenance ensures that the operation of the facilities is smooth and safe. Coupled with the maintenance to be effective, estimating repair maintenance cost is important for facility managers to identify required maintenance tasks and allocate budgets to keep facilities functional and safe. Hence, the importance of calculating repair and maintenance cost is becoming more significant to advance apartment complex facilities management. More specifically, accurate repair maintenance cost calculation has a substantial impact on budgeting and management. Proper costing helps facility managers avoid unexpected cost increases or budget overruns and allow budgets to be allocated resourcefully. In addition, accurate costing helps them determine and prioritize which maintenance tasks need to be performed. Furthermore, from the perspective of the long-term asset management, well-predicted repair and maintenance costs is directly linked with improved comfort and satisfaction of residents by maintaining the facilities in good condition and expediting any necessary repairs. For these reasons, accurate and efficient cost estimation is a key to optimize the operation and maintenance of facilities and increase the satisfaction of residents.

However, developing models and methods to predict repair and maintenance costs for apartment complexes is difficult and challenging due to various factors and their interactions affecting the costs, such as age of buildings, facility types and conditions, local climatic conditions, property values affected by market trends and so on (Kim *et al.*, 2019). Especially, a lack of large quantity of high-quality data sets is another issue to develop prediction models or methods. In addition, consistent and reliable data is required to secure the reliability of cost prediction models. Data consistency and accuracy are directly related to model performance (Ali *et al.*, 2010). Even if there is the availability of such data, it still remains questionable whether the irregularity and project-specific data could produce well-developed prediction models (Choi and Lee, 2016). In addition, repair and maintenance costs often have a non-linear relationship, and contain unexpected uncertainties. For example, sudden facility failures or urgent repairs are hard to predict. Capturing and handling these nonlinearities and uncertainties in models is thought-provoking (Lee *et al.*, 2014).

To fill these gaps, this study is aims to propose a new modeling framework that can systematically and scientifically predict the repair and maintenance costs of apartment buildings, by employing empirical data. To this end, this study develops a deep learning-driven repair and maintenance cost prediction model for apartment complexes. The significance of this study is drawn as the scalability and extensibility of the proposed model, which can be applied to various building groups in the future and contribute to the advancement of FM.

2. Literature review: knowledge gaps

2.1 Repair and maintenance cost predictions for facilities management: are they focused on apartments?

FM plays a key role in preserving and extending the life of physical assets, such as buildings, equipment and infrastructure (Bröchner *et al.*, 2019). It is known that FM by building types is dependent on the operational efficiency, safety and security,

maintenance, energy efficiency and so on [Atkin and Brooks \(2021\)](#). This is because each of the buildings has a specific purpose and operating requirements. In terms of operational efficiency, FM tailors to the type of buildings to improve smooth operation and customer satisfaction. On the aspect of safety and security, necessary preventive and countermeasures can be taken according to the type of buildings, and the buildings and end-users can be protected from potential hazards. Maintenance management extends the life of a building, while helping prevent unexpected failure or damage. In addition, energy efficiency can optimize a building's energy use to reduce energy costs and promote environmental sustainability. Hence, proper FM along with the consideration of building types ensures the efficiency and safety of building operation and contributes to cost reduction and sustainable operation ([Atkin and Brooks, 2021](#)).

Given the significance of FM, many research efforts were made in terms of repair and maintenance cost prediction for various types of built assets, such as offices, industrial buildings, university buildings and civil infrastructure. For example, [Au-Yong et al. \(2014\)](#) proposed a model to predict the maintenance costs of office buildings using condition-based maintenance. The model utilizes the building's condition to plan maintenance activities and predict costs, aiming to optimize maintenance operations. [Krstić and Marenjak \(2017\)](#) presented a model to estimate the maintenance and operation costs of university buildings. The model accounts for the specific characteristics of university buildings to predict and optimize the costs associated with maintenance and operations. [Ghodoosi et al. \(2018\)](#) focused on optimizing maintenance costs for bridge structures by employing system reliability analysis and genetic algorithms; their approach aims to find a balance between structural reliability and maintenance costs. [Meshref et al. \(2022\)](#) implemented a deep learning prediction model for the life cycle costs of industrial buildings based on different building structure alternatives; their model utilizes deep learning techniques to predict life cycle costs and assist in making optimal decisions regarding building structures. [Bayzid et al. \(2016\)](#) offered a case study on predicting maintenance costs for road construction equipment, introducing a model that considers various factors to predict maintenance costs that aids in the effective planning of maintenance activities for road construction equipment.

It is acknowledged that apartments are a residential form of multiple households living together. Hence, the frequency and impact of repair and maintenance are directly related to the living environments of tenants. However, to the best of the authors' knowledge, only a few studies focused on repair and maintenance costs with a very limited number of the relevant activities, such as roof waterproofing construction, elevator, painting and roof repairs ([Lee and Chae, 2015](#)) or with building exteriors and outdoor facilities in apartments ([Lee and Chae, 2016](#)). Although many efforts on repair and maintenance cost prediction studies were made, there is very little known about those for apartments in the FM context. To fill this knowledge gap, actual repair and maintenance cost records are collected from a real-world apartment complex and classified into seven different cases where repair and maintenance costs were spent.

2.2 Deep learning algorithm-driven approaches for improved facilities management

Given the importance of FM, research has recently been actively conducted in various sectors to advance it ([Jeong et al., 2020](#)). Moreover, there is a change in the environment of the expansion of the facility operation and maintenance market, due to the recent expansion of large-scale facilities, and a change in perspective that the management and operation of facility assets should be viewed together with facility operation and maintenance ([Jeong et al., 2020](#)). In this sense, regular maintenance and timely repairs are drawn as a key to prevent deterioration and ensure a long service life, while providing safe and secured living environments to residents and visitors.

As facilities become more advanced, larger and more complex, the need for facility asset management is increasing for reasons such as the increase in various demands from various stakeholders, while the demand for introducing AI technology for convergence with smart technologies, such as the IoT and the sensors necessary for this, is also increasing (Bröchner *et al.*, 2019). To meet these demands, many deep learning studies for facility management have been conducted. Sanzana *et al.* (2022) discussed the use of deep learning in facility management and maintenance for HVAC systems. In detail, they introduced a deep learning model termed Convolutional Neural Network (CNN) to analyze sensor data, such as temperature, humidity and airflow to monitor facility conditions and detect anomalies. The research results demonstrate that deep learning-based facilities management and maintenance methods can enhance the accuracy and efficiency of management and methods. Wang *et al.* (2020) studied the prediction of building thermal load using both shallow machine learning and deep learning techniques. They proposed using shallow machine learning algorithms and deep learning algorithms, such as Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN), to predict the thermal loads in buildings. The research results indicate that deep learning models outperform shallow machine learning models in terms of accuracy and provide more accurate thermal load predictions. Huang *et al.* (2021) introduced a hybrid deep neural network model for short-term electricity price forecasting. This combines Long Short-Term Memory (LSTM) and MLP to consider long-term dependencies and temporary fluctuations. The research results demonstrate that the proposed model achieves superior performance in electricity price forecasting, providing accurate and reliable price predictions to participants in the electricity market. Bouktif *et al.* (2020) presented a multi-sequence LSTM–RNN deep learning model combined with metaheuristics for electric load forecasting; their model utilizes the LSTM–RNN architecture and incorporates metaheuristic algorithms to enhance forecasting accuracy. The research results show that the proposed model outperforms other methods in terms of electric load forecasting, indicating its effectiveness in predicting electricity demand. Xu *et al.* (2019) offered a novel deep learning method to improve the prediction performance of indoor temperature in public buildings. They propose a deep learning model that combines CNN and LSTM to capture spatial and temporal dependencies in the data. The research results demonstrate that the proposed deep learning method achieves better prediction performance for indoor temperature, enhancing comfort and energy efficiency in public buildings.

As seen above, research on facilities management utilizing deep learning algorithms is focused on occupancy detection or energy management, while research on repair and maintenance cost is deficient. Moreover, considering the data characteristics of apartment complex repair and maintenance costs, this study proposed an LSTM model with excellent ability to process time series data. To confirm this, RNN and GRU-based models were also used simultaneously. LSTM, RNN and GRU can all capture the order and dependency of time series data well, which makes them suitable for time series data, such as the maintenance costs of an apartment complex. Furthermore, these models are effective in learning long-term dependencies and could simultaneously reflect short-term changes and long-term trends in apartment complex repair and maintenance costs.

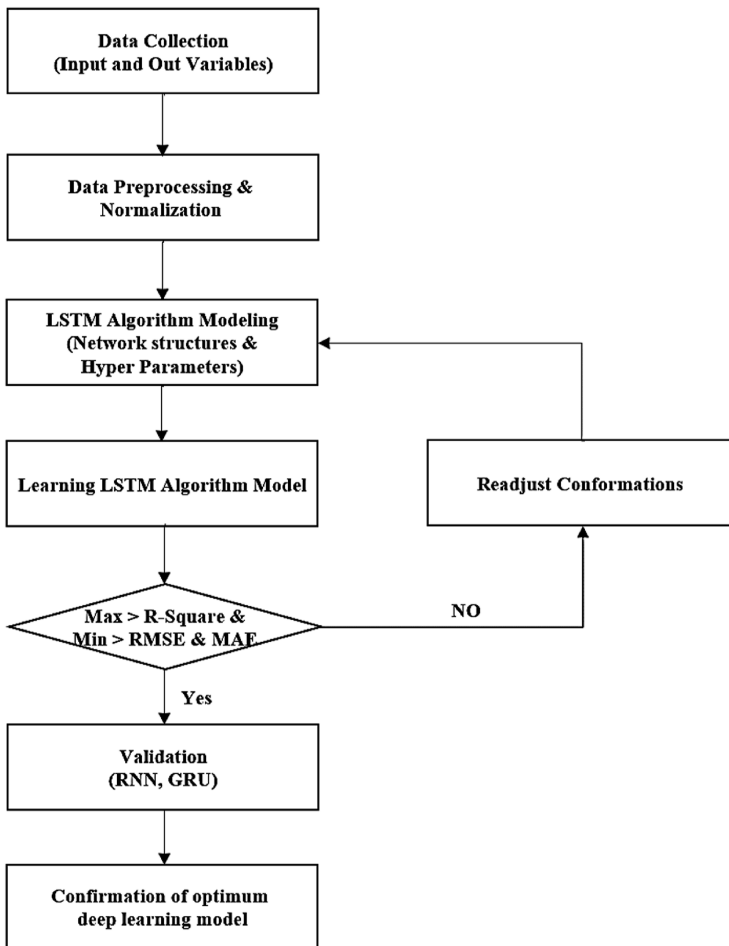
By undertaking these advantages, this study proposes an LSTM-based learning model and compares it using RNN and GRU to validate whether the proposed framework can effectively predict the repair and maintenance costs of apartment complexes. The models considered in this study are expected to improve the prediction accuracy of repair and maintenance costs by considering various characteristics of apartment complexes. Ultimately, this study aims to provide a highly reliable prediction model that can respond to various scenarios to help apartment complex managers plan resources more efficiently, and manage the budget required for repair and maintenance more effectively.

3. Research objective and methods

The main objective of this study is to propose a modeling framework to develop repair and maintenance cost prediction model based on the LSTM algorithm, by applying repair and maintenance cost data generated in an actual large-scale apartment complex. The specific procedure of this study is as follows: First, data on repair and maintenance costs incurred in actual apartment complexes is collected, along with various input variables. Second, the LSTM algorithm model is taught by dividing the collected data into input and output variables. Third, two other models (RNN and GRU) are taught using the same data. Fourth, the *R*-square, mean absolute error (MAE) and root mean square error (RMSE) for each model are computed and used to compare and verify the proposed LSTM-driven model. Figure 1 illustrates the procedure of the proposed modeling framework.

The followings are assumptions and limitations of this study:

- (1) It is noteworthy that in general, depending on types of building structures and materials (e.g. wood, steel, reinforced concrete), maintenance costs would differ from



Source(s): Author's own work

Figure 1.
Proposed modeling
framework

each other. This study confines to reinforced concrete structures, considering that typical building structures of apartment complexes in South Korea are constructed using reinforced concrete.

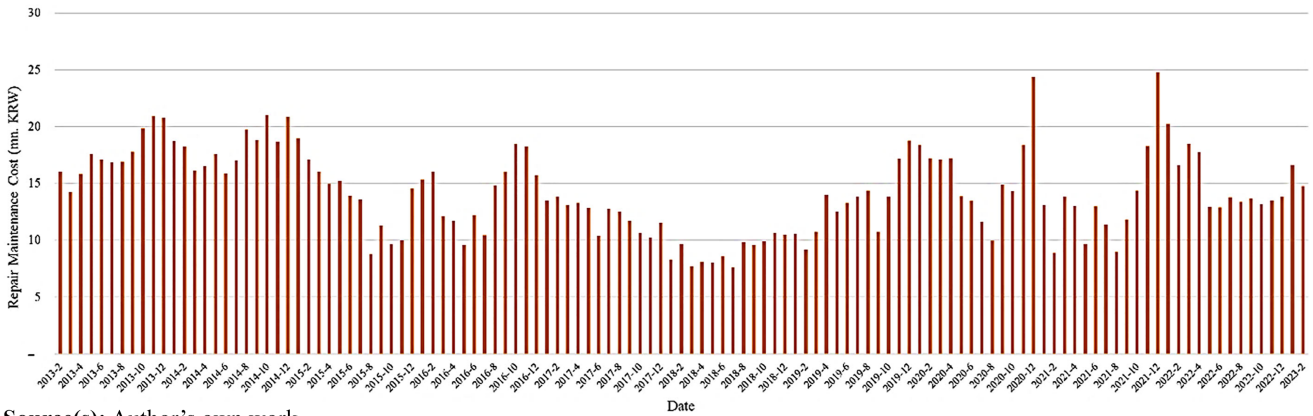
- (2) Repair and maintenance costs of buildings are different from each other, depending on building functions or types. Hence, there might be tons of different factors affecting the repair and maintenance costs. Against the possibility of such general aspects, the scope of this study was confined to apartment complexes already built. The variables used for deep learning were achieved from real-world repair and maintenance records of the project site used in this study, by categorizing repair and maintenance activities into different cases where repair and maintenance costs were spent. Given the full consideration of repair and maintenance records adopted, it was assumed that these are enough to address repair and maintenance cost predictions. However, it should be noted that more consideration of various types of factors affecting those costs would be part of future research. Nevertheless, it is assumed that external factors, especially meteorological aspects, were sufficiently considered, which are directly tied with repair and maintenance of apartment complexes.

4. Data collection

To develop a repair and maintenance cost prediction model for apartment complexes based on the LSTM algorithm, repair and maintenance cost data from a large-scale apartment complex was collected. The apartment complex used for data collection is an apartment complex with 1,774 households located in Seongnam City, South Korea. This apartment complex, completed in 1994, consists of a total of 22 buildings, each building ranging (10–25) floors. This apartment complex was selected because it is a large-scale residential complex and is an old apartment complex that is almost 30 years old, so it is possible to secure long-term repair and maintenance cost data. The collected data is the repair and maintenance costs from 2013 to 2023. This expenditure history included information such as the detailed history of repair costs, and the month in which they were incurred. [Figure 2](#) shows the monthly distribution of collected maintenance costs. There are small and large changes in the monthly repair and maintenance costs. Moreover, the explanatory power of the model was increased by utilizing weather indicators that affect the repair and maintenance costs.

Weather has a significant correlation with FM of buildings and facilities ([Krausmann et al., 2019](#); [Sulaiman et al., 2020](#)). To improve the reliability of the proposed framework, by incorporating differences in the frequency of occurrence of severe weather and weather conditions by region into the model, appropriate preparedness and preventive measures can be planned for each region, which can contribute to efficient cost management, while helping secure building safety ([Kim et al., 2021](#)). To this end, basic meteorological datasets were collected from the nearest meteorological station (i.e. City of Seongnam) to the apartment complex.

[Table 1](#) describes each of the variables. The output variable was the repair and maintenance cost, while the input variables were calendar month, usage type, building age, maximum temperature, minimum temperature, highest daily precipitation, maximum instantaneous wind speed, average humidity and total solar radiation. Those variables were adopted through a solid review of previous studies ([Au-Yong et al., 2014](#); [Bayzid et al., 2016](#); [Bröchner et al., 2019](#); [Choi and Lee, 2016](#); [Kim et al., 2021](#); [Lee et al., 2014](#); [Qiao et al., 2022](#); [Sanzana et al., 2022](#); [Zhou and Song, 2020](#)). In general, monthly repair and maintenance costs are often affected by changes in seasonal or long-term patterns. For example, the costs associated with heating systems may increase during winter season, while cooling systems during summer. To incorporate these real-world situations into the modeling process,



Source(s): Author's own work

Figure 2.
Monthly repair
maintenance cost

Variables	Description	Category
Output	Repair and maintenance cost	Numeral (KRW)
Input	Calendar month	Nominal (1–12)
	Usage type	Nominal (1–7)
	1. Building exterior area	
	2. Building interior area	
	3. Amenities such as power supply, fire suppression, lift and smart home network infrastructure	
	4. Water, gas, waste disposal and air circulation provisions	
	5. Boiler and hot water arrangements	
	6. Additional outdoor amenities and welfare facilities outside the building	
	7. Other building maintenance	
	Building age	Numeral
	Maximum temperature	Numeral (°C)
	Minimum temperature	Numeral (°C)
	Highest daily precipitation	Numeral (mm)
	Maximum instantaneous wind speed	Numeral (m/s)
	Average humidity	Numeral (%)
	Total solar radiation	Numeral (MJ/m ²)

Source(s): Author's own work

Table 1.
Description of variables

calendar month variables were included, while identifying maintenance trends that repeat in a certain month.

In addition to the temporal aspect, two major building characteristics that can affect repair and maintenance costs of apartment complexes: (1) repair and maintenance costs where those are spent (i.e. usage type variable in this study) and (2) building ages. When it comes to the usage types, a total of 7 different cases were applied, such as building exterior or interior areas, amenities, water, gas, waste disposal and air circulation provisions, boiler and hot water arrangements, additional outdoor amenities and welfare facilities outside the building and others. This is because maintenance costs come directly from various types of maintenance needs. For example, residential apartments may require maintenance of amenities, while commercial apartments may require maintenance of common areas and commercial facilities. In addition, it is acknowledged well that the age of building represents increasing needs of maintenance due to structural damages and aging equipment installed in the building.

Those building characteristics associated with repair and maintenance costs are closely linked with weather conditions that affect the level of operating heating, ventilation, air conditioning systems or cause structural damage. In this sense, six different monthly weather datasets were collected from the Korea Meteorological Administration and used as input variables, including maximum and minimum temperatures, highest daily precipitation, maximum instantaneous wind speed, average humidity and total solar radiation. In detail, high

temperatures often cause thermal expansion of building exteriors and structural materials, repeating expansion and contraction that lead to cracks and damage to building materials. In addition, high temperatures during summer affect increasing operation time of cooling systems and thus increase the frequency of maintenance. Furthermore, during summer, excessive solar radiation can increase cooling loads, resulting in more maintenance needs. On the other hand, low temperatures often incur repair costs to replace frozen pipes, while increasing operation time of heating systems that need more chances to be repaired.

Some other meteorological factors are often related to the quality of indoor conditions and structural safety. For example, higher precipitation levels cause leaks in roofs or exterior walls or flooding in underground space, which significantly increase repair and maintenance costs. Sometimes, rapid increases in rainfall make drainage systems overloaded, which requires to resolve such damages or blockages of the systems. In addition, strong winds cause damage to building exterior materials (e.g. windows, exterior wall panels) and roof structures, and this damage requires immediate repair work, significantly impacting maintenance costs. Such hazards caused by strong winds (e.g. falling objects, damaged exterior materials) require emergency maintenance, which may result in additional costs. Furthermore, humidity would promote mold growth, which reduces the quality of the indoor condition and causes health problems, so mold removal and moisture management work requires additional maintenance costs. Additionally, high humidity can promote rotting of wood and other building materials, which can lead to structural damage and costly repairs.

Table 2 summarizes descriptive statistics for the input and output variables. Figure 3 shows that the distribution was confirmed by visualizing the input and output variables with a histogram.

5. Data processing

To gain insights into the repair and maintenance cost data for apartment complexes, a sophisticated data model leveraging LSTM was devised, and subjected to verification using RNN and GRU models. The collected dataset spans from years 2013–2023, encapsulating a comprehensive timeline of repair and maintenance cost track records. These costs tend to accrue gradually over time, exhibiting temporary patterns and trends that may influence future repair and maintenance expenditures. It is noteworthy that repair and maintenance costs at a specific point in time might deviate from historical costs, signifying a presence of long-term dependencies within the dataset. Considering that the relationship between repair and maintenance costs and their future counterparts may be linear, the development of predictive models necessitates the incorporation of mechanisms that are capable of modeling nonlinear relationships (Abumohsen *et al.*, 2023; Chitalia *et al.*, 2020; Fan *et al.*, 2019).

Variables	Sample	Min	Max	Mean	Std. Deviation
Repair and maintenance cost	3,444	0.00	20100974.00	2099725.03	2473745.36
Calendar month	3,444	1.00	12.00	6.46	3.47
Usage type	3,444	1.00	7.00	4.00	2.00
Building Age	3,444	19.00	30.00	23.65	2.94
Maximum temperature	3,444	8.70	39.30	24.79	8.66
Minimum temperature	3,444	−18.60	21.70	2.18	11.34
Highest daily precipitation	3,444	0.40	285.00	40.08	42.85
Max. instant.* wind speed	3,444	9.10	24.00	13.02	2.43
Average humidity	3,444	53.00	89.00	69.48	7.64
Total solar radiation	3,444	153.76	731.25	397.91	131.19

Note(s): *Maximum instantaneous

Source(s): Author's own work

Table 2.
Descriptive statistics
of variables

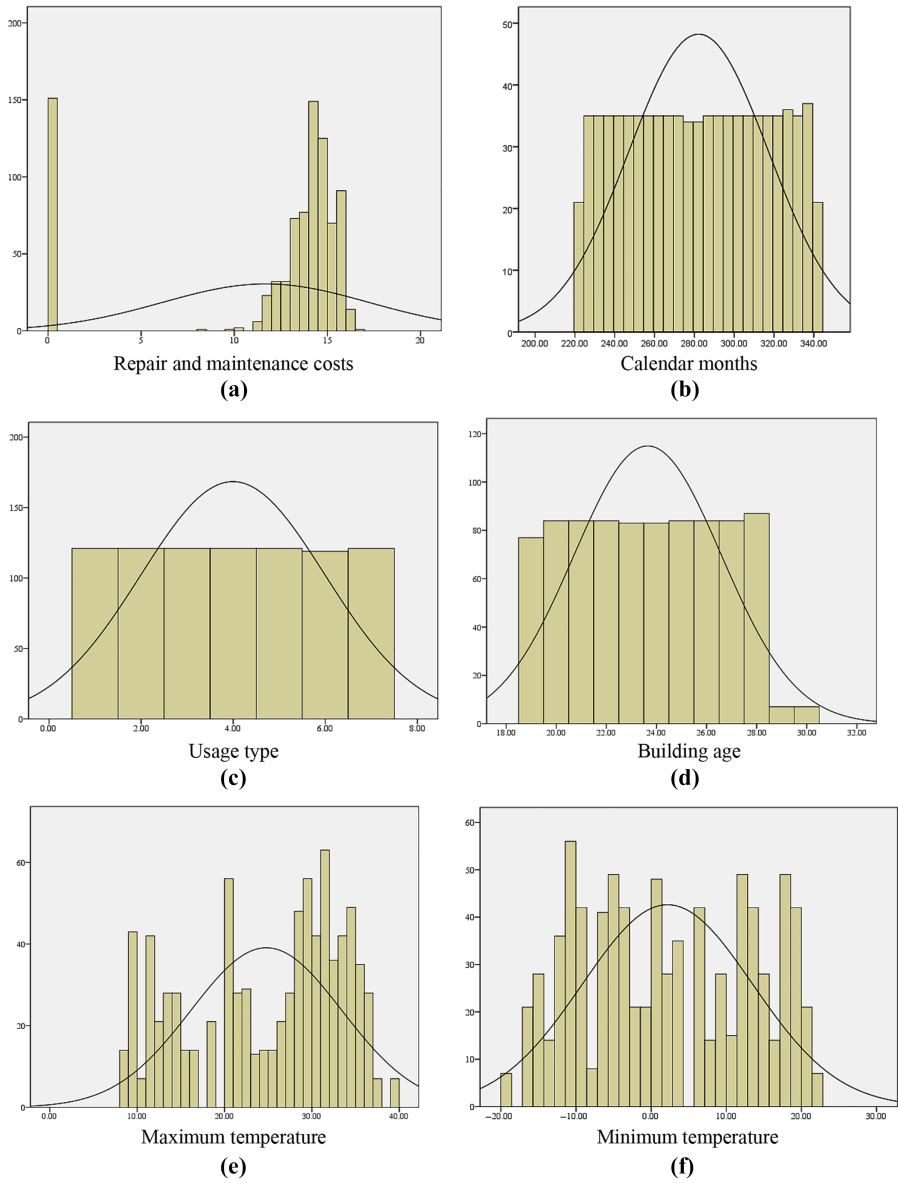
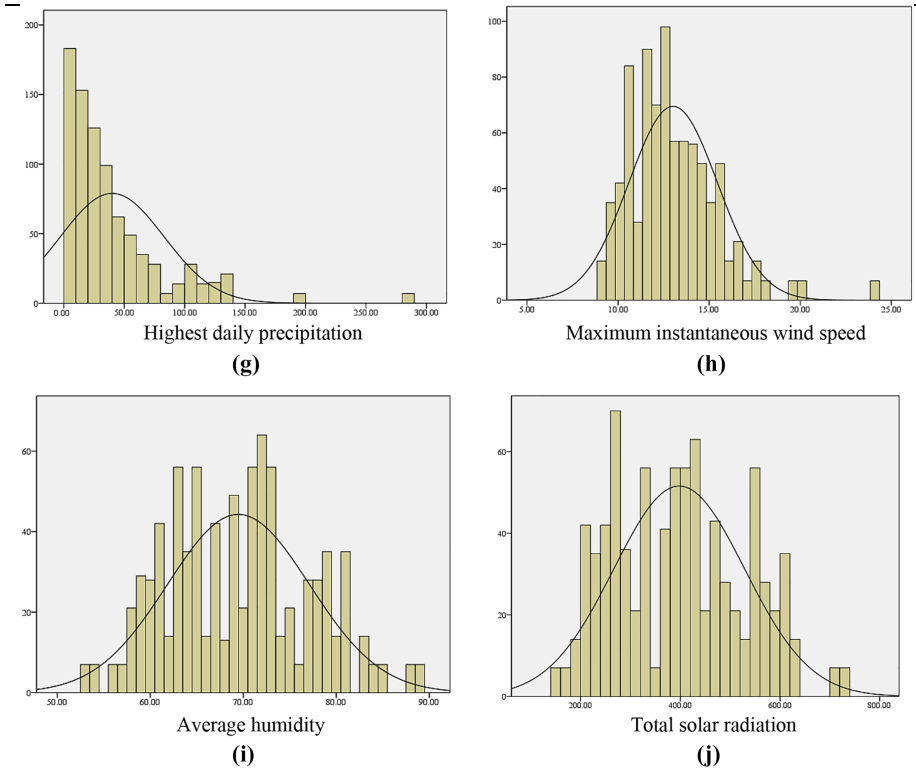


Figure 3.
Distribution of
variables

(continued)



Source(s): Author's own work

The LSTM, a variant designed to address the Vanishing Gradient problem commonly encountered in RNNs, introduces memory cells and a gating mechanism, enabling the selective retention or forgetting of information as time progresses. The robustness of LSTM in capturing extended data dependencies makes it particularly well-suited for effective modeling of the intricate long-term sequence of repair costs employed in this study (He *et al.*, 2019). Consequently, the LSTM algorithm was employed to train on the dataset, and the resultant learned outcomes were meticulously assessed through metrics, such as the R -squared value, MAE and RMSE. R -squared stands out as a conventional metric for evaluating artificial neural network models, offering insights into how much of the variance in the dependent variable is elucidated by the independent variable. Meanwhile, MAE, representing the absolute average of the residuals between actual and predicted values, serves as a crucial indicator, with higher MAE values suggesting more substantial prediction errors. The RMSE, as the square root of the mean square of the residuals, becomes particularly pertinent for gauging the magnitude of prediction errors, with larger RMSE values indicating more pronounced discrepancies.

For the input data, pre-processing was involved to normalize the data distribution using the z -score normalization method, associated with the transformation through log conversion. A strategic division was made, allocating 70% of the meticulously preprocessed data for training purposes, with 30% of the training dataset designated for validation. The remaining 30% of the total data was reserved as a distinct testing dataset.

6. Learning modeling using LSTM algorithm

It is well known that as depicted in Figure 4, LSTMs undertake 3 inputs at a time, which consists of the current input, the previous hidden state and the previous cell state and go through 3 gate processes including forget, input and output gates (Abumohsen *et al.*, 2023; Bouktif *et al.*, 2020).

In detail, the forget gate (f_t) is computed by a sigmoid activation function, which aims to produce a vector of values between 0 and 1. This process determines which information would be discarded or retained from the previous cell state (c_{t-1}). By taking the input gate (i_t) and candidate cell state (g_t) together, it decides which new information would be transferred to the new cell state (c_t). The input gate (i_t) aims to determine which information (x_t) is related to add from the current time step. The candidate cell state (g_t) is a new set of values that could be incorporated into the new cell state, which is operated with the hyperbolic tangent activation function to control the network for improved model performance. When updating the previous cell state (c_{t-1}) to the new cell state (c_t), by multiplying the previous cell state with the forget gate, certain information from the previous state could be forgotten. Then, by performing a point-wise addition with the new and relevant information achieved from the input gate and candidate cell state, the cell state could be updated properly. Lastly, the output gate (o_t) determines which information the hidden state (h_t) should carry over and then a \hat{y} prediction is computed using the current hidden state. These sequential operation processing allows the LSTM to determine which information should be retained and forgotten.

In the realm of LSTM, the convergence of network structure and hyperparameters exerts a profound influence on the predictive model, underscoring the critical importance of identifying the optimal network structure and hyperparameters (Xia *et al.*, 2023; Yang *et al.*, 2017). Hyperparameters encompass factors such as activation size, batch placement, optimizer, epoch determination and more, each of which being contingent upon the specific network structure scenario. Given the myriad potential combinations of network structures and hyperparameters, discerning which elements warrant attention becomes paramount (Zhou and Song, 2020). Moreover, the process of pinpointing these optimal components holds considerable significance, steering the development of a predictive model that is highly efficient, while also being capable of effectively capturing the complex patterns and dependencies that lie within the data. These refinements play a pivotal role in enhancing both the predictive capabilities and the overall performance of the model. Furthermore, the

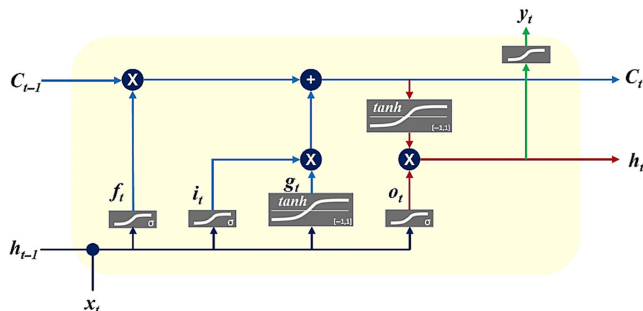


Figure 4.
LSTM modeling
process

Source(s): Author's own work

manipulation of network structure and hyperparameters, derived through a meticulous optimization process, contributes significantly to elevating the generalization capacity of deep learning models. Employing a simulation-based tuning approach allows for a nuanced exploration of various combinations of network structures and hyperparameters, empowering the model to navigate through diverse data settings and real-world scenarios with enhanced predictive prowess and robustness.

Activation functions play a crucial role in enabling models to learn and represent nonlinear relationships. In this study, the Rectified Linear Unit (ReLU) was selected as the activation function. ReLU overcomes the limitations of the sigmoid function, by passing the input value, if it is greater than 0, unchanged; and if it is less than 0, outputting 0. These characteristics have established ReLU as an actively utilized activation function in various deep learning applications (Lim and Kim, 2023). The batch size, determining the number of data samples in each training step, significantly influences the learning rate. Therefore, maintaining an appropriate learning rate is crucial to stabilizing the learning process. The epoch indicates how many times the model learns the training data through repetition, representing the process of feeding the entire training data set to the model at once (Lim and Kim, 2023). In this study, the model was effectively trained on data using 1,000 epochs and a batch size of 5. The optimizer is an algorithm that minimizes the loss function by adjusting weights, controlling the model's stability and managing the learning rate. Hence, selecting an appropriate optimizer is essential to enhancing the performance of deep learning models (Krizhevsky *et al.*, 2012).

Notably, generally used optimizers, such as RMSprop, AdaGrad, Adam and Adadelta, possess unique characteristics, which influence the model's learning process differently. RMSprop adjusts the learning rate to uniformly regulate the update rate of each parameter (Huk, 2020). In contrast, AdaGrad automatically adjusts the learning rate of parameters based on the update history (Qiao *et al.*, 2022). Adadelta stabilizes the optimization process by adjusting the learning rate using only recently updated information (Sester *et al.*, 2018). Adam is an adaptive learning rate algorithm that can adjust the learning rate for each parameter (Zheng *et al.*, 2019). Due to the distinct features and limitations of each optimizer, selecting the most suitable one for a specific model or the data characteristics is crucial. In this study, RMSprop, AdaGrad, Adadelta and Adam were individually simulated, and compared to identify the optimal optimizer.

Table 3 presents the R -squared value, MAE and RMSE as the learning results for each model, based on the optimizer and the number of layers. The optimal model was chosen based on the highest R -squared value, lowest MAE and RMSE. As indicated in Table 3, the

Algorithm	Optimizer	Number of layer	R -Square	MAE	RMSE
LSTM	RMSprop	1	0.7251	0.3008	0.5102
		2	0.7661	0.2284	0.4609
		3	0.8261	0.1907	0.3803
	AdaGrad	1	0.2539	0.6035	0.8573
		2	0.2544	0.5539	0.8547
		3	0.0561	0.5789	0.9671
	Adadelta	1	-0.5571	0.7948	1.0631
		2	-0.1328	0.7111	1.0473
		3	-0.1210	0.6293	1.0471
	Adam	1	0.4071	0.4657	0.7344
		2	-0.1033	0.5548	1.0183
		3	-0.0997	0.5567	1.0281

Source(s): Author's own work

Table 3.
Learning outcomes of
the LSTM model

best performance was achieved when using RMSprop as the optimizer and incorporating three layers ($R^2 = 0.8261$; MAE = 0.1907; RMSE = 0.3803).

7. The robustness validation of model

7.1 Setting compared learning outcomes using RNN and GRU models

To validate the robustness of the developed LSTM algorithm-driven deep learning model, this study constructed RNN and GRU models. RNNs are well-suited for sequential modeling, particularly in tasks like time series prediction, as they can process data sequences, and capture temporal dependencies through internal memory (Fekri *et al.*, 2021). GRUs represent a variant of RNNs designed to address the Vanishing Gradient problem, offering a balance between complexity and efficiency when capturing temporal dependencies. The GRU provides a simplified architecture compared to LSTM, featuring fewer gating mechanisms (Pham *et al.*, 2023). Given the advantages, both RNN and GRU models were developed, and their learning performances by holding the modeling scenario that was applied to the LSTM modeling procedure (i.e. four different types of optimizers, three different cases of the number of layer). Accordingly, Table 4 summarizes the learning outcomes of RNN and GRU models. The learning outcomes reveal that both models excelled with an RMSprop optimizer and three hidden layers.

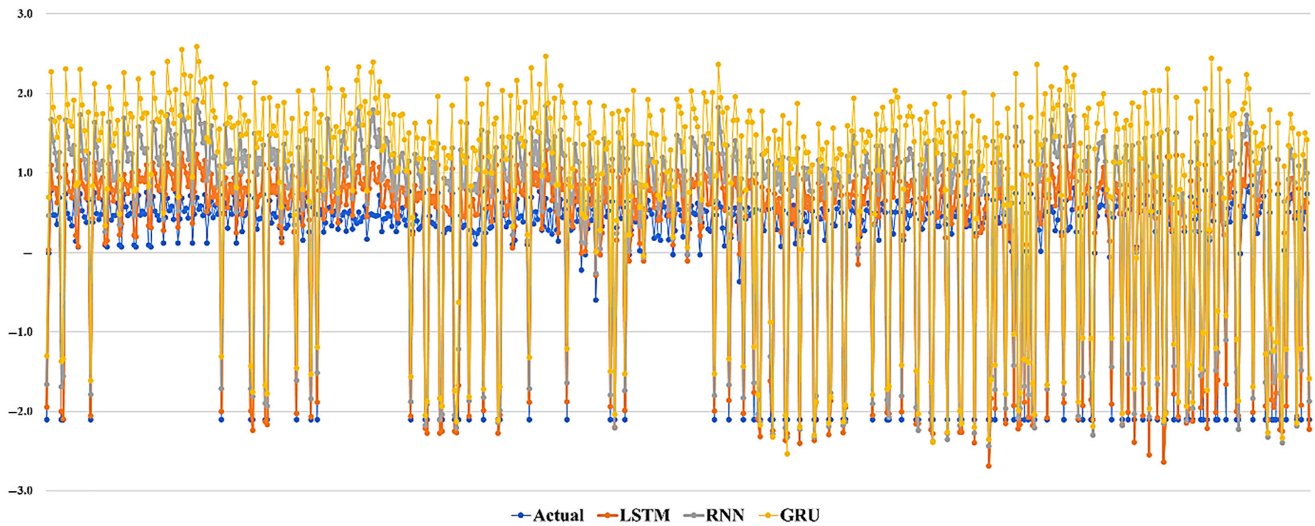
7.2 Predicted-to-actual comparison analysis: LSTM versus RNN and GRU models

To scientifically assess the reliability and accuracy of the developed LSTM model and the compared models (i.e. RNN, GRU), a predicted-to-actual comparison analysis was conducted. As depicted in Figure 5, the predicted results closely aligned with the actual values,

Algorithm	Optimizer	Number of layer	R-Square	MAE	RMSE
RNN	RMSprop	1	0.7156	0.3201	0.5205
		2	0.7347	0.2433	0.5084
		3	0.8107	0.1915	0.3957
	AdaGrad	1	0.2804	0.5745	0.8188
		2	0.0789	0.5507	0.8968
		3	0.0399	0.5633	0.9690
	Adadelta	1	-0.5254	0.8692	1.1325
		2	-0.2708	0.7176	1.1055
		3	-0.3429	0.7228	1.1360
	Adam	1	0.0352	0.5464	0.9711
		2	0.2486	0.4852	0.8652
		3	-0.0610	0.5516	1.0012
GRU	RMSprop	1	0.7064	0.3053	0.5103
		2	0.7834	0.2208	0.4502
		3	0.7985	0.2034	0.4099
	AdaGrad	1	0.1878	0.5967	0.8836
		2	0.1086	0.6027	0.9361
		3	0.1174	0.5506	0.9233
	Adadelta	1	-0.3255	0.8303	1.1087
		2	-0.0731	0.6916	1.0167
		3	-0.1075	0.6476	1.0282
	Adam	1	-0.2071	0.5945	1.0630
		2	-0.2071	0.5911	1.0890
		3	-0.1660	0.5807	1.0754

Table 4. Learning outcomes of the RNN and GRU models

Source(s): Author's own work



Source(s): Author's own work

Figure 5.
Learning results of the
optimization model

confirming a high level of reliability and accuracy. However, when having a closer look into the learning performances of LSTM, RNN and GRU based on *R*-squared value, MAE and RMSE, the LSTM model holds the highest *R*-squared value of 0.8261, compared to 0.8107 by RNN and 0.7985 by GRU. Similarly, the results of MAE and RMSE of the LSTM model indicates the lowest error values of 0.1907 and 0.3803, respectively. These results confirm that the LSTM model’s learning is more accurate and reliable, compared to the RNN and GRU models’ learning performance.

Based on the proven robustness of the LSTM model, Table 5 summarizes the composition of the final model developed in this study. As the post-processing, this study scrutinized the prediction results of the validation and test datasets to identify potential overfitting issues in the model. Overfitting occurs when a model overly conforms to training data, hindering its generalization to new data, and compromising predictive performance and reliability. As shown in Table 6, for the validation and test datasets, the MAE was found to be approximately 0.204 and 0.226, while the RMSE was 0.313 and 0.423, respectively. These results indicate that the disparity in results from the validation dataset was not solely due to model overfitting, but rather substantiated the model’s validity for testing on a new dataset.

The validation results of the robustness confirm algorithmic characteristics of LSTM, RNN and GRU. RNNs are a type of network that processes sequential data by remembering information achieved from the previous inputs. LSTM and GRU are known as special types of RNN, which aims to address the problem of vanishing gradients in RNNs. This problem occurs when the gradients of the weights are small to learn. Compared to typical RNN, LSTM is designed to selectively store and forget information for long periods. On the other hand, GRU is a simplified version of LSTM, which is easier and faster to learn, but not effective at storing and accessing for long periods. Given the nature of three different algorithms, the modeling results achieved from this study represent well such algorithmic characteristics and support the learning performance of LSTM.

8. Discussion

This study proposed a predictive framework for anticipating the repair and maintenance costs in apartment complexes using the LSTM algorithm. The research involved gathering real-world data on repair and maintenance expenditures in apartment complexes and categorizing

Table 5.
Final model profile

Arrangement Algorithm		Specifics LSTM
Hyper-parameter	Optimizer	RMSprop
	Batch Size	5
	Epoch	1,000
	Activation Function	ReLU
Network structure	Layer	3
	Node	48–64–80
	Source(s): Author’s own work	

Table 6.
Comparison of results
across data sets

Data set	MAE	RMSE
Validation	0.204	0.313
Test	0.226	0.423
Source(s): Author’s own work		

various influential variables. To improve the learning process of the LSTM algorithm model, simulations were conducted, incorporating adjustments to hyperparameters and network scenarios, followed by a comprehensive analysis of the outcomes. In addition, to assess the model's efficiency, RNN and GRU algorithms were applied to the same dataset, enabling a comparative analysis and the validation of each model's results. The research findings highlight the exceptional performance of the LSTM model, especially when employing the RMSprop optimizer with three layers, surpassing the performance of other models. A validation process was implemented to ensure the reliability of the ultimately selected LSTM model, confirming a high degree of similarity between the predicted results and the actual values within the learning dataset. Furthermore, a thorough comparison of the predicted results from both the validation and test datasets was conducted to identify potential overfitting issues, and to verify the model's reliability and generalizability.

The adaptability of this model extends across diverse sectors. For example, construction firms might utilize LSTM models to predict repair and maintenance costs during housing construction, assisting in the development of construction budgets and maintenance plans. Insurance companies might leverage LSTM models for a precise evaluation of repair costs, streamlining the insurance claims process. Also, real estate agents and developers might incorporate the model into the budgeting process for repairs and renovations, supplying clients with dependable and precise budget estimates. Additionally, governmental agencies might effectively manage repair and maintenance budgets for urban infrastructure by employing LSTM models. Strategic planning and budgeting for repair and maintenance work might be guided by the anticipated costs. The integration of advanced devices and technologies, such as smart homes in apartment facilities, allows for synergistic collaboration with the LSTM model, opening avenues for innovative foundational research. In the public sector, LSTM models function as valuable tools for budgeting the repair and maintenance of urban infrastructure, including bridges and roads. Reliance on estimated costs can significantly improve the efficiency of planning and budgeting for repair and maintenance tasks. Housing management organizations might also embrace LSTM models to streamline budget management and renovations for housing complexes or apartment facilities, optimizing budget allocation through anticipating repair costs in advance.

This model demonstrates versatility in predicting maintenance costs across a spectrum of industries. In the aviation sector, the LSTM model can be employed to anticipate and estimate repair and maintenance expenses for aircraft, empowering airlines to effectively plan and allocate budgets based on expected costs. Similarly, within the automotive industry, the LSTM model can project repair and maintenance costs for vehicles, fostering precise sharing of cost information between service centers and vehicle owners, thereby enhancing efficient budget management. In the energy industry, the LSTM model's application extends to predicting costs related to the repair and maintenance of power plants and facilities. This would enable energy companies to proficiently oversee maintenance budgets and uphold system reliability. Moreover, the LSTM model finds applicability across various building types, including industrial facilities, public structures and commercial buildings, for the prediction and management of maintenance costs. Consequently, the LSTM model holds significant value in anticipating maintenance costs across a variety of industries and building types, playing a pivotal role in guiding budget allocation and planning.

Nevertheless, conducting further research is encouraged by involving more supplementary data from diverse apartment complexes, cities and countries to improve the quality of the proposed modeling framework. Given the study's concentration on a specific building type, additional research encompassing data collection, comparison and analysis for various building types is essential to expand the model's applicability.

9. Conclusion

In recent times, there has been a noticeable surge in the complexity of buildings, primarily driven by trends such as the proliferation of high-rise structures, and the construction of larger buildings. As a result, there is escalating demand for highly sophisticated organizational structures and intricate management strategies to skillfully navigate and oversee these evolving developments. This increased complexity underscores the growing importance of advanced FM, as the intricacies of modern buildings necessitate more sophisticated and nuanced approaches to maintenance and operation. Furthermore, the field of FM is undergoing a transformation due to a variety of factors that include strengthened regulations, evolving standards, rapid technological advances and increasing emphasis on sustainability. These factors highlight the significance of adopting comprehensive and systematic FM practices to effectively address the multifaceted challenges presented by modern structures. The imperative for advanced FM is further emphasized by the dynamic nature of such external factors, demanding adaptive and proactive approaches to ensure the efficient functioning of buildings and their components.

Especially, achieving advanced FM pivots on a critical component, involving the precise allocation and meticulous planning of budgets. The accuracy of these financial considerations is of paramount importance, especially in the context of predicting high-frequency repair and maintenance costs. The intricacies of modern buildings call for a predictive model that is capable of anticipating these costs with precision. The ability to make accurate predictions empowers facility managers to effectively estimate and allocate budgetary resources, enabling informed decisions regarding investment priorities and the development of proactive maintenance strategies to address potential issues before they escalate. Furthermore, the optimization of resource allocation and prioritizing of maintenance activities are integral aspects of advanced facility management. By these means, unnecessary budgetary waste can be curtailed, and the extension of asset lifespan through the implementation of suitable maintenance technologies becomes feasible. The integration of accurate predictions for repair and maintenance costs further facilitates cost-efficiency analyses, allowing facility managers to fine-tune and optimize their maintenance investments in accord with overarching strategic objectives. The precise calculation of repair and maintenance costs plays a pivotal role in supporting decision-making processes related to asset lifecycle management. This involves maximizing the value of assets by strategically planning for maintenance and repairs. The accurate prediction of costs provides facility managers with the necessary insights to make informed decisions that contribute to effective asset management practices. In essence, accurate forecasting of repair and maintenance costs is indispensable for advanced facility management, serving as a linchpin for effective budgeting, facilitating the implementation of preventive maintenance strategies, conducting thorough cost-efficiency analyses and adopting optimized asset management practices. As the complexity of buildings continues to evolve, the role of accurate predictions becomes increasingly vital in navigating the challenges and intricacies of modern FM.

This study proposed the development of a new model based on deep learning algorithms that utilizes authentic repair and maintenance cost data obtained from apartment complexes. The envisioned model aims to predict economic losses associated with facility maintenance in apartment complexes with a high degree of effectiveness. The proposed model provides a valuable tool that can contribute to mitigating financial management risks and reducing losses in forthcoming apartment construction projects.

Despite the significance of this study, it should be still noted that repair and maintenance costs of many different facilities are different from each other, depending on building functions or types. In other words, as there might be tons of different factors affecting the repair and maintenance costs, more consideration of various types of factors

affecting those costs would be part of future research. Nevertheless, it is noteworthy that the modeling framework approach proposed in this study is capable of managing many different sets of variables affecting repair and maintenance costs of certain facilities and thus to achieve more reliable and accurate predicted values. The main findings and framework established through this study hold the potential for broad applicability, as they can be extended to various building types and diverse industries within the expansive field of FM.

References

- Abumohsen, M., Owda, A.Y. and Owda, M. (2023), "Electrical load forecasting using LSTM, GRU, and RNN algorithms", *Energies*, Vol. 16 No. 5, p. 2283, doi: [10.3390/en16052283](https://doi.org/10.3390/en16052283).
- Ali, A.S., Kamaruzzaman, S.N., Sulaiman, R. and Cheong Peng, Y. (2010), "Factors affecting housing maintenance cost in Malaysia", *Journal of Facilities Management*, Vol. 8 No. 4, pp. 285-298, doi: [10.1108/14725961011078990](https://doi.org/10.1108/14725961011078990).
- Atkin, B. and Brooks, A. (2021), *Total Facility Management*, Wiley-Blackwell, New York.
- Au-Yong, C.P., Ali, A.S. and AhmAd, F. (2014), "Prediction cost maintenance model of office building based on condition-based maintenance", *Eksploratacja i Niezawodnosc-Maintenance and Reliability*, Vol. 16 No. 2, pp. 319-324.
- Bayzid, S.M., Mohamed, Y. and Al-Hussein, M. (2016), "Prediction of maintenance cost for road construction equipment: a case study", *Canadian Journal of Civil Engineering*, Vol. 43 No. 5, pp. 480-492, doi: [10.1139/cjce-2014-0500](https://doi.org/10.1139/cjce-2014-0500).
- Bouktif, S., Fiaz, A., Ouni, A. and Serhani, M.A. (2020), "Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting", *Energies*, Vol. 13 No. 2, p. 391, doi: [10.3390/en13020391](https://doi.org/10.3390/en13020391).
- Bröchner, J., Haugen, T. and Lindkvist, C. (2019), "Shaping tomorrow's facilities management", *Facilities*, Vol. 37 Nos 7/8, pp. 366-380, doi: [10.1108/f-10-2018-0126](https://doi.org/10.1108/f-10-2018-0126).
- Chitalia, G., Pipattanasomporn, M., Garg, V. and Rahman, S. (2020), "Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks", *Applied Energy*, Vol. 278, 115410, doi: [10.1016/j.apenergy.2020.115410](https://doi.org/10.1016/j.apenergy.2020.115410).
- Choi, S.-M. and Lee, Y.-S. (2016), "Economic analysis for setting appropriate repair cycles on the fixed materials and facilities in the public rental housing", *Advances in Materials Science and Engineering*, Vol. 2016, pp. 1-10, doi: [10.1155/2016/7423801](https://doi.org/10.1155/2016/7423801).
- Fan, C., Wang, J., Gang, W. and Li, S. (2019), "Assessment of deep recurrent neural network-based strategies for short-term building energy predictions", *Applied Energy*, Vol. 236, pp. 700-710, doi: [10.1016/j.apenergy.2018.12.004](https://doi.org/10.1016/j.apenergy.2018.12.004).
- Fekri, M.N., Patel, H., Grolinger, K. and Sharma, V. (2021), "Deep learning for load forecasting with smart meter data: online Adaptive Recurrent Neural Network", *Applied Energy*, Vol. 282, 116177, doi: [10.1016/j.apenergy.2020.116177](https://doi.org/10.1016/j.apenergy.2020.116177).
- Ghodoosi, F., Abu-Samra, S., Zeynalian, M. and Zayed, T. (2018), "Maintenance cost optimization for bridge structures using system reliability analysis and genetic algorithms", *Journal of Construction Engineering and Management*, Vol. 144 No. 2, 04017116, doi: [10.1061/\(asce\)co.1943-7862.0001435](https://doi.org/10.1061/(asce)co.1943-7862.0001435).
- He, F., Zhou, J., Feng, Z.-k., Liu, G. and Yang, Y. (2019), "A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm", *Applied Energy*, Vol. 237, pp. 103-116, doi: [10.1016/j.apenergy.2019.01.055](https://doi.org/10.1016/j.apenergy.2019.01.055).
- Huang, C.J., Shen, Y., Chen, Y.H. and Chen, H.C. (2021), "A novel hybrid deep neural network model for short-term electricity price forecasting", *International Journal of Energy Research*, Vol. 45 No. 2, pp. 2511-2532, doi: [10.1002/er.5945](https://doi.org/10.1002/er.5945).

- Huk, M. (2020), "Stochastic optimization of contextual neural networks with RMSprop", *Intelligent Information and Database Systems: 12th Asian Conference, ACIIDS 2020, Phuket, Thailand, March 23-26, 2020, Proceedings, Part II 12*, pp. 343-352, doi: [10.1007/978-3-030-42058-1_29](https://doi.org/10.1007/978-3-030-42058-1_29).
- Islam, R., Nazifa, T.H., Mohammed, S.F., Zishan, M.A., Yusof, Z.M. and Mong, S.G. (2021), "Impacts of design deficiencies on maintenance cost of high-rise residential buildings and mitigation measures", *Journal of Building Engineering*, Vol. 39, 102215, doi: [10.1016/j.jobe.2021.102215](https://doi.org/10.1016/j.jobe.2021.102215).
- Jeong, Y., Lee, Y. and Jung, Y. (2020), "Requirements management framework for design management and information characteristics", *Korean Journal of Construction Engineering and Management*, Vol. 21 No. 6, pp. 3-15.
- Kim, S., Lee, S. and Han Ahn, Y. (2019), "Evaluating housing maintenance costs with loss-distribution approach in South Korean apartment housing", *Journal of Management in Engineering*, Vol. 35 No. 2, 04018062, doi: [10.1061/\(asce\)me.1943-5479.0000672](https://doi.org/10.1061/(asce)me.1943-5479.0000672).
- Kim, J.-M., Yum, S.-G., Park, H. and Bae, J. (2021), "A deep learning algorithm-driven approach to predicting repair costs associated with natural disaster indicators: the case of accommodation facilities", *Journal of Building Engineering*, Vol. 42, 103098, doi: [10.1016/j.jobe.2021.103098](https://doi.org/10.1016/j.jobe.2021.103098).
- Krausmann, E., Girgin, S. and Necci, A. (2019), "Natural hazard impacts on industry and critical infrastructure: natech risk drivers and risk management performance indicators", *International Journal of Disaster Risk Reduction*, Vol. 40, 101163, doi: [10.1016/j.ijdrr.2019.101163](https://doi.org/10.1016/j.ijdrr.2019.101163).
- Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012), "Imagine classification with deep convolutional neural networks", *Advances in Neural Information Processing Systems*, Vol. 25.
- Krstić, H. and Marenjak, S. (2017), "Maintenance and operation costs model for university buildings", *Tehnički Vjesnik*, Vol. 24 No. 1, pp. 193-200.
- Lee, K. and Chae, C. (2015), "The characteristics of the repair cost distribution in apartment housing", *KIEAE Journal*, Vol. 15 No. 6, pp. 19-26, doi: [10.12813/kieae.2015.15.6.019](https://doi.org/10.12813/kieae.2015.15.6.019).
- Lee, K.-H. and Chae, C.-U. (2016), "Repair cost estimation model of the building exterior and outdoor facilities in apartment housing", *KIEAE Journal*, Vol. 16 No. 3, pp. 129-135, doi: [10.12813/kieae.2016.16.3.129](https://doi.org/10.12813/kieae.2016.16.3.129).
- Lee, K., Ahn, Y. and Chae, C.-U. (2014), "Establishment of the repair cycle of the components of the apartment housing", *KIEAE Journal*, Vol. 14 No. 2, pp. 69-75, doi: [10.12813/kieae.2014.14.2.069](https://doi.org/10.12813/kieae.2014.14.2.069).
- Lim, K.-K. and Kim, J.-M. (2023), "Financial loss assessment for weather-induced railway accidents based on a deep learning technique using weather indicators", *Applied Sciences*, Vol. 13 No. 18, 10418, doi: [10.3390/app131810418](https://doi.org/10.3390/app131810418).
- Meshref, A., El-Dash, K., Basiouny, M. and El-Hadidi, O. (2022), "Implementation of a life cycle cost deep learning prediction model based on building structure alternatives for industrial buildings", *Buildings*, Vol. 12 No. 5, p. 502, doi: [10.3390/buildings12050502](https://doi.org/10.3390/buildings12050502).
- Pham, T., Le-Hong, T. and Tran, X. (2023), "Efficient estimation and optimization of building costs using machine learning", *International Journal of Construction Management*, Vol. 23 No. 5, pp. 909-921, doi: [10.1080/15623599.2021.1943630](https://doi.org/10.1080/15623599.2021.1943630).
- Qiao, R., Wang, R. and Zou, Y. (2022), "A brief analysis on damaged building classification: optimizer and learning rate", *International Conference on Cloud Computing, Performance Computing, and Deep Learning (CCPCDL 2022)*.
- Sanzana, M.R., Maul, T., Wong, J.Y., Abdulrazic, M.O.M. and Yip, C.-C. (2022), "Application of deep learning in facility management and maintenance for heating, ventilation, and air conditioning", *Automation in Construction*, Vol. 141, 104445, doi: [10.1016/j.autcon.2022.104445](https://doi.org/10.1016/j.autcon.2022.104445).
- Sester, M., Feng, Y. and Thiemann, F. (2018), "Building generalization using deep learning", *The ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 4 No. 42, pp. 565-572, doi: [10.5194/isprs-archives-xlii-4-565-2018](https://doi.org/10.5194/isprs-archives-xlii-4-565-2018).

-
- Sulaiman, N., Abid, S.K., Nazir, U., Mahmud, N.P.N., Latib, S., Hafidz, H., Roslan, A.F. and Rahim, N. (2020), "Need for resilience healthcare facilities management (RHFM) in Malaysia's public hospitals. A critical literature review", Proceedings of International Conference Industry Engineering Operation Management.
- Wang, Z., Hong, T. and Piette, M.A. (2020), "Building thermal load prediction through shallow machine learning and deep learning", *Applied Energy*, Vol. 263, 114683, doi: [10.1016/j.apenergy.2020.114683](https://doi.org/10.1016/j.apenergy.2020.114683).
- Xia, Y., Wang, J., Wei, D. and Zhang, Z. (2023), "Combined framework based on data preprocessing and multi-objective optimizer for electricity load forecasting", *Engineering Applications of Artificial Intelligence*, Vol. 119, 105776, doi: [10.1016/j.engappai.2022.105776](https://doi.org/10.1016/j.engappai.2022.105776).
- Xu, C., Chen, H., Wang, J., Guo, Y. and Yuan, Y. (2019), "Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method", *Building and Environment*, Vol. 148, pp. 128-135, doi: [10.1016/j.buildenv.2018.10.062](https://doi.org/10.1016/j.buildenv.2018.10.062).
- Yang, C., Trudel, E. and Liu, Y. (2017), "Machine learning-based methods for analyzing grade crossing safety", *Cluster Computing*, Vol. 20 No. 2, pp. 1625-1635, doi: [10.1007/s10586-016-0714-2](https://doi.org/10.1007/s10586-016-0714-2).
- Zheng, Z., Lu, P. and Pan, D. (2019), "Predicting highway-rail grade crossing collision risk by neural network systems", *Journal of Transportation Engineering, Part A: Systems*, Vol. 145 No. 8, 04019033, doi: [10.1061/jtpebs.0000257](https://doi.org/10.1061/jtpebs.0000257).
- Zhou, S. and Song, W. (2020), "Deep learning-based roadway crack classification using laser-scanned range images: a comparative study on hyperparameter selection", *Automation in Construction*, Vol. 114, 103171, doi: [10.1016/j.autcon.2020.103171](https://doi.org/10.1016/j.autcon.2020.103171).

Corresponding author

Junseo Bae can be contacted at: junseobae@korea.ac.kr

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com