

Article

Not peer-reviewed version

ANN-Based Real-Time Prediction of Heat and Mass Transfer in the Paper-Based Storage Enclosure for Sustainable Preventive Conservation

[Bo Han](#) , [Fan Wang](#) ^{*} , [Julie Bon](#) , [Linda Mac Millan](#) , [Nicholas K Taylor](#)

Posted Date: 20 May 2025

doi: [10.20944/preprints202505.1551.v1](https://doi.org/10.20944/preprints202505.1551.v1)

Keywords: heat and mass transfer; artificial neural network (ANN); prediction, hygrothermal condition; buffering effect; storage enclosure; preventive conservation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

ANN-Based Real-Time Prediction of Heat and Mass Transfer in the Paper-Based Storage Enclosure for Sustainable Preventive Conservation

Bo Han ¹, Fan Wang ^{1, *}, Julie Bon ², Linda MacMillan ² and Nick K Taylor ³

¹ School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh EH14 4AS, UK

² National Library of Scotland, Edinburgh EH1 1EW, UK

³ School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK

* Correspondence: fan.wang@hw.ac.uk

Abstract: The storage enclosures are vital for stabilising the micro-environment within, facilitating preventive conservation efforts and enabling energy savings by reducing the need for extensive macro-environmental control within the room. However, real-time conformity monitoring of the micro-environment to ensure compliance with preventive conservation specifications poses a practical challenge due to a limitation in implementing physical sensors for each enclosure. This study aims to address this challenge by using an ANN-based prediction for temperature and RH changes in response to macro-environmental fluctuations. A numerical model was developed to simulate transient heat and mass transfer between macro and micro environments, and then employed to determine an acceptable macro-environmental range for sustainable preventive conservation and to generate a dataset to train a sequence-to-sequence ANN model. This model was specially designed for real-time prediction of heat and mass transfer and to simulate the micro condition under varying levels of control accuracy over the macro environment. The effectiveness of the prediction model was tested through a real trial application in the laboratory revealed a robust prediction of micro-environments inside different enclosures under various macro-environmental conditions. This modelling approach offers a promising solution for monitoring the micro-environmental conformity and further implementing the relaxing control strategy in the macro-environment without compromising the integrity of the collections stored inside the enclosures.

Keywords: heat and mass transfer; artificial neural network (ANN); prediction; hygrothermal condition; buffering effect; storage enclosure; preventive conservation

1. Introduction

Preventive conservation of heritage collections plays an important role in minimising climate-induced decay and avoiding detrimental effect when they are maintained in an indoor environment as stable as possible [1]. As many of these collections are made of organic materials such as paper which shrink and expand with dropping and rising moisture content, especially fluctuations in relative humidity (RH) should be avoided to prevent mechanical damage at a fast strain rate in the storage space of libraries, galleries, and museums [2]. Climate change exacerbates this risk, accelerating degradation, increasing conservation costs, and necessitating climatization [3]. While maintaining a stable environment is essential, it requires precision air conditioning with tight control, resulting in substantial energy consumption and carbon emissions [4,5]. Effective conservation and protection of heritage collections require well-informed management tools grounded in an understanding of how environmental parameters evolve over time and vary across the environment [6]. In addition, a great number of collections are usually stored in some European historic buildings with poor energy performance, posing a challenge for retrofitting the buildings where exterior

façades are decorated with carvings and sculptures while interior walls are decorated with frescoes [7]. Consequently, improving operational energy efficiency is a challenging task for achieving sustainable preventive conservation.

Many approaches have been developed for controlling the micro-environment within enclosures rather than whole macro-environment in the room. Adding some buffering materials inside the enclosures to achieve passive control is a prevalent approach [8]. The enclosure separates the micro-environment inside from the macro-environment, the room where the enclosure presents and the confined space with buffering materials provide a degree of hygrothermal buffering effect.

In the National Library of Scotland (NLS), the paper-based enclosures are used for storing books in storage rooms, like many other cultural institutions. These enclosures are made of functional cardboard with specific elements and porous structure. The cardboard consists of three layers: an external water-proof layer, a middle multi-thin layer, and an internal lining layer. The external layer contains polypropylene to provide water-proof and thermal-insulation functions [9,10]. The middle and inner layers contain pulp cotton and calcium carbonate to regulate the micro-environmental temperature and humidity [11,12]. To consider fabric permeability, a thin multi-layer structure is designed in the middle layer to increase the tortuosity of porous channels. This design can increase moisture retention through small air gaps between these layers, thus improving the buffering effect [13].

The enclosures can stabilize the micro-environment, especially mitigating fluctuations in RH due to its buffering effect. This buffering effect has been rigorously evaluated in a laboratory test [14]. It quantified the enclosure's capacity to minimize the micro-environmental fluctuations over some given periods under two distinct macro-environmental conditions (10%RH fluctuation within 50%~60%RH at fixed 20°C for humidity buffering test; 5°C fluctuation within 17~22°C at fixed 50%RH for temperature buffering test). The results demonstrated that the micro-environmental fluctuations are 2% to 8%RH and 1 to 2°C smaller than the macro-environmental ones. These findings confirmed that the enclosures can provide a degree of buffering capacity to mitigate micro-environmental fluctuation, by reducing its response to that of the macro-environment. Such buffering capacity enables a relaxing control for macro-environment when the micro-environment complies with the standards of preventive conservation. This observation aligns with a key insight from an Italian study [15], suggesting that the control accuracy of precision air conditioning systems can be relaxed within the macro-environment, while still ensuring effective preventive conservation within the micro-environment. However, achieving real-time conformity monitoring of the micro-environment to ensure compliance with preventive conservation specifications is constrained by the practical difficulty of implementing multitude physical sensors inside the enclosures to map their conditions.

Numerical model that simulates heat and mass transfer between macro and micro environments can predict the micro-environment in response of the macro-environment. This technique has been proven effective for conducting hygrothermal analysis [16]. It was used to develop a model to simulate coupled diffusive heat and mass transfer between hygroscopic building structure and indoor air. The model was used to investigate the indoor humidity and air quality in a timber house after its validation with a set of field measurement data [17]. Also based on this technique, a full dynamic hygrothermal model was developed to couple heat and moisture transfer in porous building materials to predict the hygrothermal performance of buildings while incorporating convection and advection [18]. These models excel in capturing hygrothermal interaction between the micro and macro environments with good representations of the physical principles. Their application not only enhances understanding of the hygrothermal interaction but also establishes a foundation for assessing the conformity of preventive conservation within the micro-environment and further achieving the relaxing control of the macro-environment.

However, the heat and mass simulation is highly time-consuming, typically running in hours, so it is not feasible to use it practically in precision air conditioning systems with a feedback control loop that operates on a scale of minutes. To address this issue, an artificial neural network (ANN) offers a feasible solution when it is trained with a purposely prepared dataset emerges as. This

approach has seen successful application in various domains of thermal science. A mathematical model of the refrigerant cycle was created to generate data regarding evaporator cooling capacity and relevant parameters such as air flow, temperature, and RH. Subsequently, this data was used for training ANN model [19]. Similarly, in a study about structure optimization of latent heat thermal storage units, computational fluid dynamic simulations were used to describe the transient thermal behaviour. Based on the simulation results, an ANN was trained to predict stored energy [20]. Additionally, ANN models trained with historical real data have been instrumental in optimising building energy use with acceptable indoor microclimate [21].

Above studies show that ANN is an effective tool for capturing the complex operating characteristics of building system. However, there are only several studies investigated the hygrothermal interaction in the realm of building science by integrating traditional mathematical modelling with ANN. Remarkably, no prior study has delved into exploring the buffering effect of the storage enclosures in facilitating a relaxing control of macro-environment for sustainable preventive conservation. Therefore, this study aims to achieve real-time conformity monitoring of micro-environment as macro-environmental control is relaxed. To achieve the aim, there are three objectives:

- to develop a coupled heat and mass transfer model for predicting the micro-environment.
- to determine acceptable macro-environment for obtaining associated range of the relaxing control.
- to train an ANN model for real-time conformity monitoring of the micro-environment.

2. Methodology

In alignment with these three stated objectives, the methodology framework comprised three parts. First, a numerical model was developed to simulate the heat and mass transfer between the two environments: the macro and micro and was validated after. Second, a three-stage method based on the heat and mass transfer simulation was employed to determine the upper limit of acceptable macro-environmental condition through a trial-and-error process. Following this, a series of simulations was run to acquire data, representing conditions under the relaxing control. In the third part, four distinct sets of data were generated within this acceptable range to capture the main characteristics of heat and mass transfer interaction between macro and micro environments under relaxed control. Leveraging these generated datasets alongside on-site collected data for tight control, a Long Short-Term Memory (LSTM) neural network was trained for real-time prediction of micro-environmental conditions. Subsequently, the robustness of this real-time prediction was assessed in a practical application. Further details of these activities are provided below.

2.1. Numerical Simulation of Heat and Mass Transfer

The model geometry was 180mm wide, 230mm tall, 80mm deep, and its envelope was 1.14mm thick, exactly those of a real enclosure used in the library (Figure 1). The laminar airflow was assumed aligns with the width of the enclosure, infiltrating through both envelope and gas of the enclosure. The air can flow through the gaps to free flow domain inside the empty enclosure and also permeate the porous domain of the cardboard.

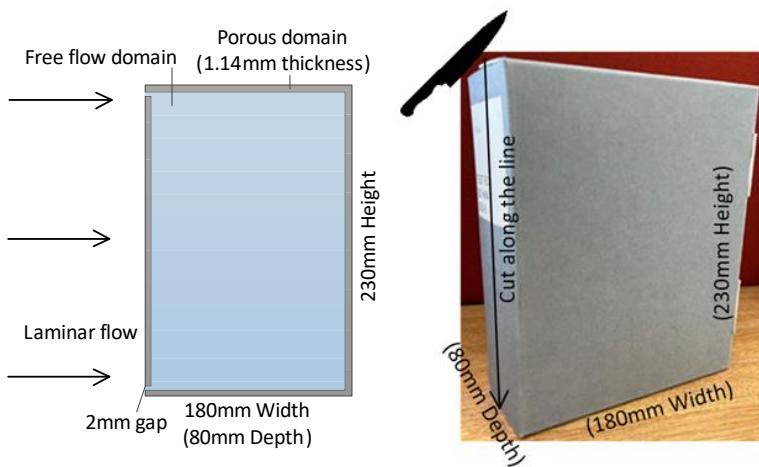


Figure 1. Geometry diagram of the model.

In the storage room, air flow velocity of supply air was about 0.2m/s (the Reynold number $= \frac{\rho v l}{\mu} = \frac{1.2 \times 0.2 \times 0.002}{1.8 \times 10^{-5}} = 26 < 2000$ laminar flow), and the temperature was between 15-25°C. The airflow was expected to affect the convective heat and mass transfer, which consequently lead to phase change and water vapor diffusion in the cardboard material. This hygrothermal response of the material was a key point to understand the buffering effect on mitigating the hygrothermal fluctuation in the micro environment inside the enclosure. The enclosure envelope could retain and release heat and moisture and its hygrothermal condition was influenced by the conditions in both the micro and macro environments. Meanwhile, the hygrothermal conditions of both the micro and macro environments were interacting with each other through both the porous material and the gap. Hence, a two-way coupling was necessary in this study to the two-fold interaction with the two hygrothermal variables.

Figure 2 shows the interdependence of heat, and moist-air transfer. The heat and moisture balance are coupled in a way that the latent heat, moisture state, transfer characteristic and thermal storage affect the heat transfer while the heat transfer determines the saturated vapor pressure, moisture driving force and storage conditions. The heat and moist-air convection balance are coupled in a way that the temperature affects the air density, and reversely the air convective heat transfer affects the temperature field. The moisture and air convection balance are coupled in a way that the air flow determines the convective moisture transfer, and in turn the vapor permeability determines the air flow in the porous media. To consider the complexity of modelling, we selected direct coupling method, regarding the water vapor and air as one fluid in both free flow domain and porous domain. The space of these two domains is continuous and a single set of conservation equations are solved by one solver in the simulation [22].

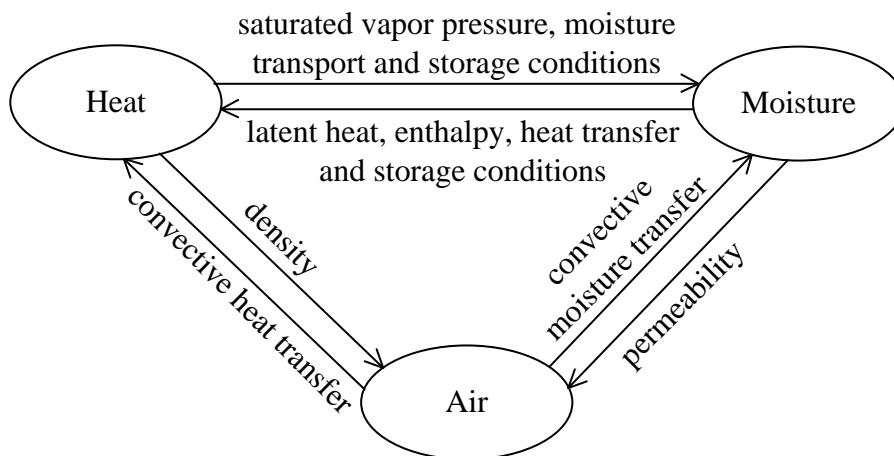


Figure 2. Interdependence of heat and moist-air transfer, which involves three components: heat, air and moisture.

The Equations 1-3 show the mathematical description [23,24]:

$$(\rho C_p)_{eff} \frac{\partial T}{\partial t} + \rho C_p u \cdot \nabla T - \nabla \cdot (k_{eff} \nabla T + L_v \delta_p \nabla (\phi p_{sat})) = Q \quad (1)$$

$$\xi \frac{\partial \phi}{\partial t} + \rho u \nabla \omega_v - \nabla \cdot (\xi D_w \nabla \phi + \delta_p \nabla (\phi p_{sat})) = G \quad (2)$$

$$\rho \left(\frac{\partial u}{\partial t} + u \cdot \nabla u \right) + \nabla \cdot (pI - K) = F \quad (3)$$

where $(\rho C_p)_{eff}$ is the effective volumetric heat capacity at constant pressure, ρ is moisture air density [kg/m^3], C_p is heat capacity [$\text{J}/(\text{kg}\cdot\text{K})$]. u is moisture air velocity [m/s]. T is the temperature [K]. k_{eff} is the effective thermal conductivity [$\text{W}/(\text{m}\cdot\text{K})$]. L_v is the latent heat of evaporation [J/kg]. δ_p is vapor permeability [s]. ϕ is the relative humidity. ω_v is vapor mass fraction. p_{sat} is the vapor saturation pressure [Pa]. Q is the heat source [$\text{W}/\text{m}^3\cdot\text{s}$]. ξ is the moisture storage capacity [kg/m^3], $\xi = \frac{\partial w(\phi_w)}{\partial \phi_w}$. D_w is the moisture diffusivity [m^2/s]. G is the moisture source [$\text{kg}/\text{m}^3\cdot\text{s}$]. p is the fluid pressure [N/m^2]. I is the identity tensor. K is the viscous force [N/m^2]. F is the external forces applied to the fluid [N/m^2].

2.1.1. Model Setting

Because of limited computing resource in transient heat and mass simulation, a 2D axial-symmetry model with depth was used to represent 3D model. The heat and mass model was developed in COMSOL. Associated settings include four parts, 1) material setting, 2) boundary and initial conditions, 3) building mesh for the model, 4) solver setting, as Table 1 shows.

Table 1. Model setting.

Description	Value	Unit
(1) Carboard material		
Density	662	kg/m^3
Thermal conductivity	0.055	$\text{W}/(\text{m}\cdot\text{K})$
Heat capacity at constant pressure	1.028	$\text{J}/(\text{kg}\cdot\text{K})$
Diffusion coefficient	1.49E-10	m^2/s
Water content	$W_c = \rho(0.1237 \times RH + 0.0034)/1000$	kg/m^3
Vapor resistance factor	95.63	-
(2) Boundary condition		
Laminar air flow	0.2	m/s
Temperature	Macro-environmental temperature	$^{\circ}\text{C}$
RH	Macro-environmental RH	%RH
Upper and lower gaps of the enclosure	open boundary	-
Initial conditions		
Temperature and RH in both domains	Macro-environmental temperature and RH at the first second	$^{\circ}\text{C}$ and %RH
Velocity in both domains	0.2	m/s
Pressure in both domains	(Ambient pressure – reference pressure)	Pa
(3) Meshing		
Element types	triangular or quadrilateral	-

No. of layers in porous media	2~4	-
Mesh density	dense in the porous domain and gradually coarse toward the centre of free flow domain	-
Maximum element growth rate	1.05	-
Maximum curvature factor	0.2	-
(4) Solver		
Time stepping	second-order BDF	-
Maximum step	0.25	h
Solving method	automatic Newton	-
tolerance factor	0.01	-
maximum No. of iterations	4	-

2.1.2. Model Validation

To meet the library's preventive conservation requirements, the macro-environment must be strictly maintained at 15-25°C and 40%-60% RH. Within these bands, different macro-environmental fluctuations affect the interaction of heat and mass transfer, especially for the hygrothermal response from porous media of the cardboard. Therefore, the measurement data used for model validation should encompass different degrees of these fluctuations.

Three sets of micro-environmental data were gathered under varying levels of macro-environmental control accuracy. The first set originated from the library's storage room, where tight control was maintained. The second and third sets were obtained in a controlled test chamber, where fluctuations of 10%RH (ranging from 50% to 60%RH) and 5°C (fluctuating between 17°C and 22°C) occurred cyclically every 60 minutes. In each case, fixed conditions of 20°C and 50%RH were maintained, respectively.

The temperature and RH in the centre of free flow domain were selected as simulated data to compare with the measured data. The maximum absolute and relative errors were used to evaluate the model accuracy [25].

$$\left\{ \begin{array}{l} \text{Max. absolute error} = \max\{|T \text{ or } RH_{\text{measurement}} - T \text{ or } RH_{\text{simulated}}|_t\} \\ \text{Max relative error} = \max\left\{\frac{(T \text{ or } RH_{\text{measurement}} - T \text{ or } RH_{\text{simulated}})}{T \text{ or } RH_{\text{measurement}}}\right\}_t \end{array} \right. \quad (4)$$

where t is the sampling time for each 15min in measurement and simulation.

Additionally, agreement between measurement and simulated data was indicated by Kling-Gupta Efficiency (KGE). It is given as Equation 5 [26].

$$KGE = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \quad (5)$$

where $\beta = \frac{\mu_s}{\mu_o}$, μ_s and μ_o are means of the simulated and measurement data with time series; $\alpha = \frac{\sigma_s}{\sigma_o}$, σ_s and σ_o are standard deviations of the simulated and measurement data with time series.

Negative KGE values are regarded as a bad model performance while the positive values show a good performance. KGE=1 indicates perfect agreement in terms of correlation, mean, and variability. KGE>0.5 indicates that a model performance is upon the baseline in empirical terms [27].

2.2. Determination of Acceptable Macro-Environment

Relaxing tight macro-environment on the site is impractical due to strict regulations of management in the storage room. Therefore, it is necessary to amplify the macro-environmental fluctuations based on the HAM simulation. The three-stage method of data acquisition was developed to generate data and further determine acceptable macro-environment. It included macro-environmental data amplification, heat and mass simulation, and acceptable macro-environment

determination (Figure 3). Stage 1 was about macro-environment data acquisition. In Stages 2 and 3, there was a loop to search the upper limit of acceptable macro-environment by using the trial-and-error method.

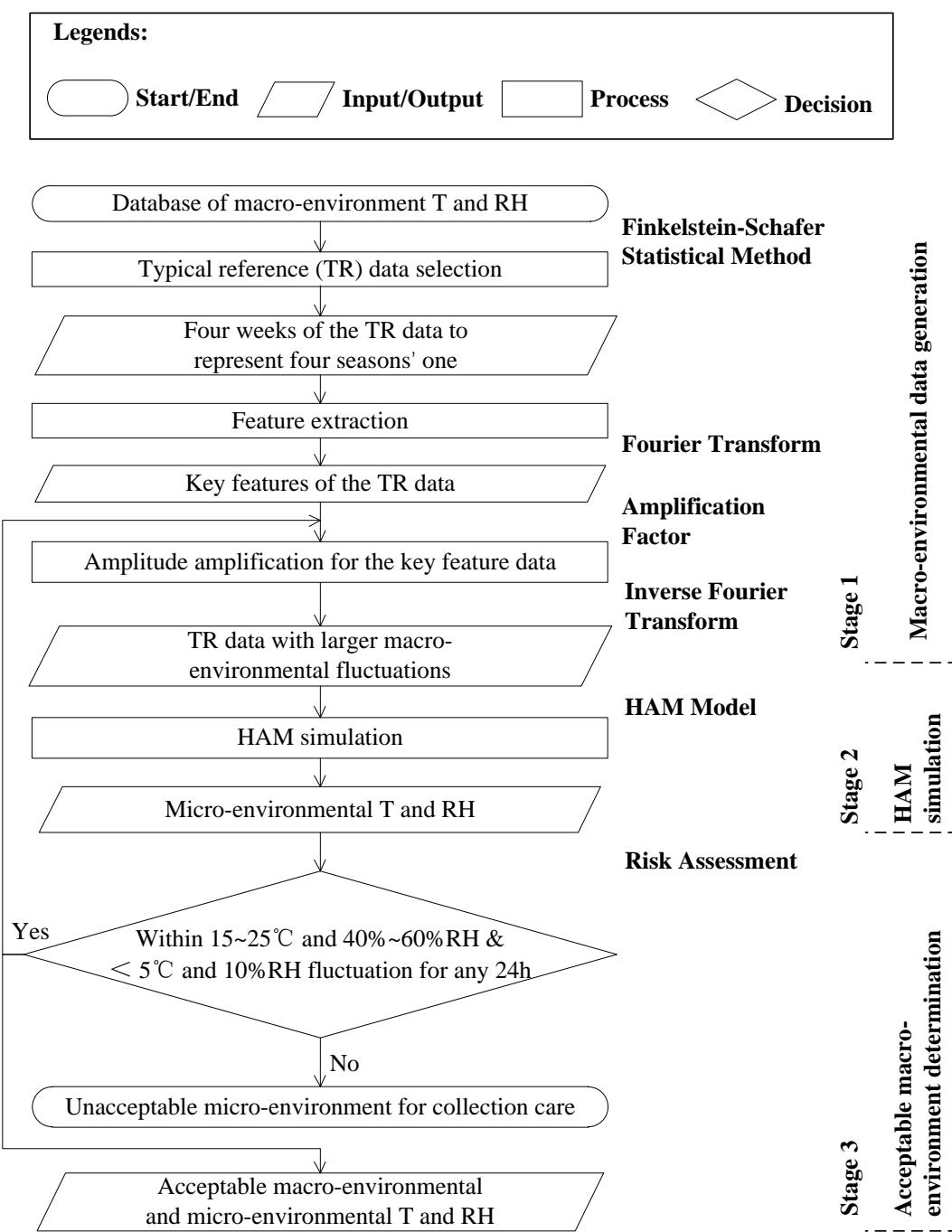


Figure 3. Framework of acceptable macro-environment determination.

2.2.1. Typical Reference Data Selection

The choice of typical reference (TR) data was intended to identify four representative weeks of data, each encapsulating the distinct characteristics of the four seasons within a two-year dataset collected from the storage room. The typical-meteorological-year (TMY) data generation method in building energy simulation was used [28]. It was based on the Finkelstein–Schafer (FS) statistical method [29]. Each week was selected within specific season in these two years. The daily means of macro-environmental temperature and RH were calculated. Then, they were sorted into an ascending

order to calculate the cumulative distribution function (Equation 6) for each week ($F_{n,w}(x)$) in long term and individual season of that week ($F_{n,w,s}(x)$) in short term.

$$F_n(x) = \begin{cases} 0, & x < x_1 \\ (k - 0.1), & x_k \leq x \leq x_{k+1} \\ 1, & x > x_n \end{cases} \quad (6)$$

where x is the daily mean of temperature or RH, k is the rank order, n is the number of x .

The FS statistic was calculated by Equation 7.

$$FS(s, w) = \sum_{k=1}^n |F_{n,w,s}(x_k) - F_{n,w}(x_k)| \quad (7)$$

where $F_{n,w,s}(x_k)$ was calculated by the daily means within that week and that season. $F_{n,w}(x_k)$ was calculated by the daily means over these two years. The typical week for each season was determined by that week which has the smallest $FS(s, w)$.

The typical reference data was amplified to obtain new data reflecting the macro-environment in the relaxing control.

2.2.2. Data Amplification

The principle of data amplification was drawn inspiration from electroencephalogram (EEG) signal processing. The EEG signal can be regarded as time-dependent amplitude function that expressed as a spectrum in different frequencies [30,31]. It was associated with five typical brain activities, including relaxed awareness without any attention (alpha waves), active thinking (beta waves), deep sleep (delta waves), deep meditation (theta waves), some certain brain diseases (gamma waves) [32]. An analogy was made between the EEG signal processing and the macro-environment data processing. The data can be decomposed into many components to represent the data features. Each key feature can be regarded as a factor which affects the macro-environmental fluctuations. The Fourier transfer (FT) is a tool to achieve data transformation between time domain (original data) and frequency domain (features). It can be expressed by Equation 8 [33].

$$X(k) = \sum_{j=1}^n x(j)W_n^{(j-1)(k-1)} \quad (8)$$

And the inverse FT can be expressed by Equation 9.

$$x(j) = \frac{1}{n} \sum_{k=1}^n X(k)W_n^{-(j-1)(k-1)} \quad (9)$$

where $X(k)$ and $x(j)$ are the FT pairs in time and frequency domains. $X(k)$ is the macro-environmental data with k points sequence. $x(j)$ can be regarded as j data features with successive amplitude at specific frequencies. $W_n = e^{(-2\pi i)/n}$ is a primitive n -th root of 1. $i = \sqrt{-1}$, the imaginary unit.

The daily RH data, $X(k)$, was decomposed to the data features, $x(j)$, by using the FT function tool in MATLAB software. Some of key features were intentionally selected to represent the original data. These key features were combined by using the inverse FT. The more key components, the more accurate representation. The number of key features was determined by a high R-squared value, $R^2 \geq 0.95$, between the original data and represented data. Sequentially, the selected key features $x(j)$ were multiplied by an amplification factor and inversely transferred to obtain the data with large fluctuation.

The trial-and-error method was used to amplify the fluctuations gradually until the micro-environment didn't comply with the specifications of preventive conservation in the risk assessment.

The data amplification of macro-environment applied to the RH only, excluding temperature, there are two reasons. First, as a useful approximation in RH control for preventive conservation, a coupled RH change of 3% occurs for each degree of temperature change [34]. A slight temperature changes possibly could cause the RH out of the control band or allowable range of 24h fluctuation. The setting of temperature change is dependent on the RH control requirement in the precision air conditioning with humidity priority because the RH change in paper-based collection storage directly impacts on paper stability and preservation. Second, the thermal mass of the library's storage room can ensure a small fluctuation for indoor temperature. Throughout the two-year dataset, the macro-environmental temperature remained consistently within the range of 17-20°C, with fluctuations of $\leq 0.2^\circ\text{C}$ occurring for 95% of the time. The relaxation of temperature control was not considered.

2.2.3. Determination Process of Acceptable Macro-Environment

In the preventive conservation, the micro-environmental data must comply with the specifications, 15~25°C, 40%~60%RH control bands and <±5°C, <±10%RH 24h-fluctuations for most paper-based collections. The risk assessment checked the micro-environmental temperature and RH which were output from the heat and mass simulation.

A sliding window calculator was developed as Equation 10 shows, which helps to check the maximum and minimum values and the biggest fluctuation in any 24h (96 sampling points). In this risk assessment, if the micro-environment complies with the specifications, the associated macro-environment can be accepted. The amplification factor was increased to amplify the macro-environmental fluctuations in the loop designed to search the upper limit of acceptable macro-environment.

$$\left\{ \begin{array}{l} \max\{RH_i, \dots, RH_{i+95}\} - \min\{RH_i, \dots, RH_{i+95}\} \leq 10\% \\ \max\{RH_i, \dots, RH_{i+95}\} \leq 60\% \\ \min\{RH_i, \dots, RH_{i+95}\} \geq 40\% \end{array} \right. \quad (10)$$

where n is the number of sliding windows, $n=2592$ (4-week of data).

2.3. ANN Modelling

2.3.1. Data Preparation

Based on above section 2.2, within the range of acceptable macro-environment, four sets of data were prepared to represent four levels of macro-environmental control accuracy in the relaxing control (Table 2).

Table 2. Four datasets in the relaxing control.

Level	24h Fluctuation (band) in macro-environment	24h Fluctuation (band) in micro-environment	Amplification factor
1	±10 (42.2~56.1) %RH	±7.1 (45.5~55) %RH	10
2	±12 (39.6~57.4) %RH	±7.9 (44.7~55.4) %RH	11.6
3	±14 (37~59.4) %RH	±8.1 (43.8~56.8) %RH	13.4
4	±16 (33~65) %RH	±9.1 (43~57.3) %RH	15

These levels from 1 to 4 signify the control accuracy of macro-environment from current control requirement level based on the specifications of preventive conservation to the most relaxing level determined in section 2.2.3. These four generated datasets, along with the on-site collected dataset in the tight control, denote varying levels of control accuracy within the acceptable conditions. They were used for ANN training. The macro-environmental temperature and RH were input, and the corresponding micro-environmental ones were output. To improve stability during the training, it requires a normalization process for all input and output data before the training [35]. An inverse normalization is necessary to predict the outputs after the training. Thus, Z-score normalization was used to rescale the data with zero mean and a unit variance here. The time-series data was divided into two parts in the training. 90% of the data was allocated for the training process and the rest of 10% was allocated for testing the trained networks to evaluate the prediction accuracy [36,37].

2.3.2. Architecture of the ANN

The real-time prediction of the micro-environment should consider the data relationship sequence-to-sequence, inputting 24h macro-environmental temperature and RH to predict the corresponding micro-environmental ones. Long short-term memory (LSTM) network, a type of ANN, was designed to solve this problem [38] (Figure 4). The architecture consisted of forget gate,

input gate and output gate to achieve appropriate-term memory. The mathematical expression about forget gate, input gate, output gate, and hidden layer are given by Equation 11 [39].

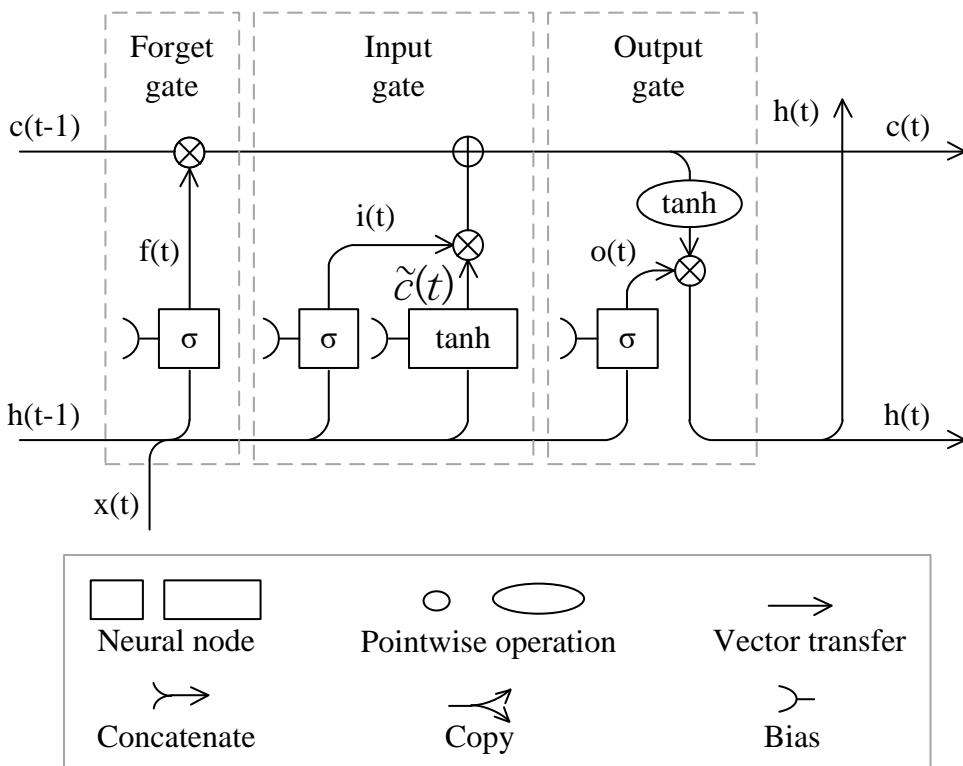


Figure 4. Architecture of LSTM memory unit.

$$\begin{cases} f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \\ i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \\ \tilde{c}_t = \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}) \\ c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \\ o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \\ h_t = o_t \cdot \tanh(c_t) \\ y_t = h_t \end{cases} \quad (11)$$

where f_t is the forget factor to decide on which information should be kept and which to forget. σ and \tanh are sigmoid (the value is between 0~1) and hyperbolic tangent (the value is between -1~1) activation functions, respectively. h_t , i_t , \tilde{c}_t , c_t and o_t denote the recurrent information, input factor, candidate cell state, cell state and output factor. W_f , W_i , $W_{\tilde{c}}$ and W_o are the weights, and b_f , b_i , $b_{\tilde{c}}$ and b_o are the bias. c_t denotes the cell state of neural network. The operator ‘.’ denotes the pointwise multiplication of two vectors. x_t and y_t are the input and output.

The LSTM network uniquely captured sequence-to-sequence relationships between macro and micro environments, demanding deep learning for intelligent decision-making during training. Its architecture included specialized layers to leverage the intelligence of its three gate-operated memory units. Dropout was a regularization technique to prevent overfitting issue in the neural network. Setting dropout layer between the hidden layers and the last hidden layer - fully connected layer (also call dense layer) could effectively prevent this issue in RNN training [40,41]. Fully-connected layer allowed the network to capture the dependence between various features by using non-linear transformations to the input data [42]. Because there was dependence between temperature and RH with highly non-linear relationship, a fully-connected layer could contribute to effective data mapping here. Regression layer was selected as output layout because the data mapping between macro and micro environments was a many-to-many regression problem.

Therefore, the basic architecture of our LSTM should include a sequence input layer, LSTM layer(s), drop layer(s), fully-connected layer, and regression output layer (Figure 5). The optimal

number of LSTM layers and associated neurons were determined by a trial-and-error method (Figure 6). We relied on past research on LSTM training to determine the maximum numbers of LSTM layers and neurons (5 and 250) for searching the optimal architecture [43,44]. To assess the accuracy of trained network, the latest RMSE and absolute errors of trained network were compared with the errors from previous training in the loop and the measurement errors ($\pm 0.5^{\circ}\text{C}$ and $\pm 3\%\text{RH}$ in the library's storage room) respectively. If they are smaller than both the earlier prediction errors and the measurement errors, it means that this trained network has minimal prediction errors and acceptable accuracy. Otherwise, the numbers of layers and neurons should be increased gradually until the best architecture was found.

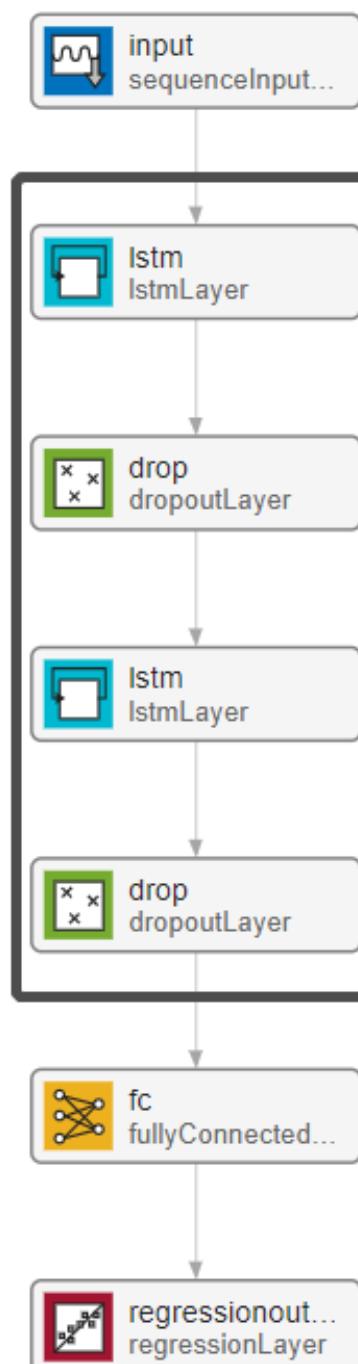


Figure 5. Basic architecture of LSTM regression network.

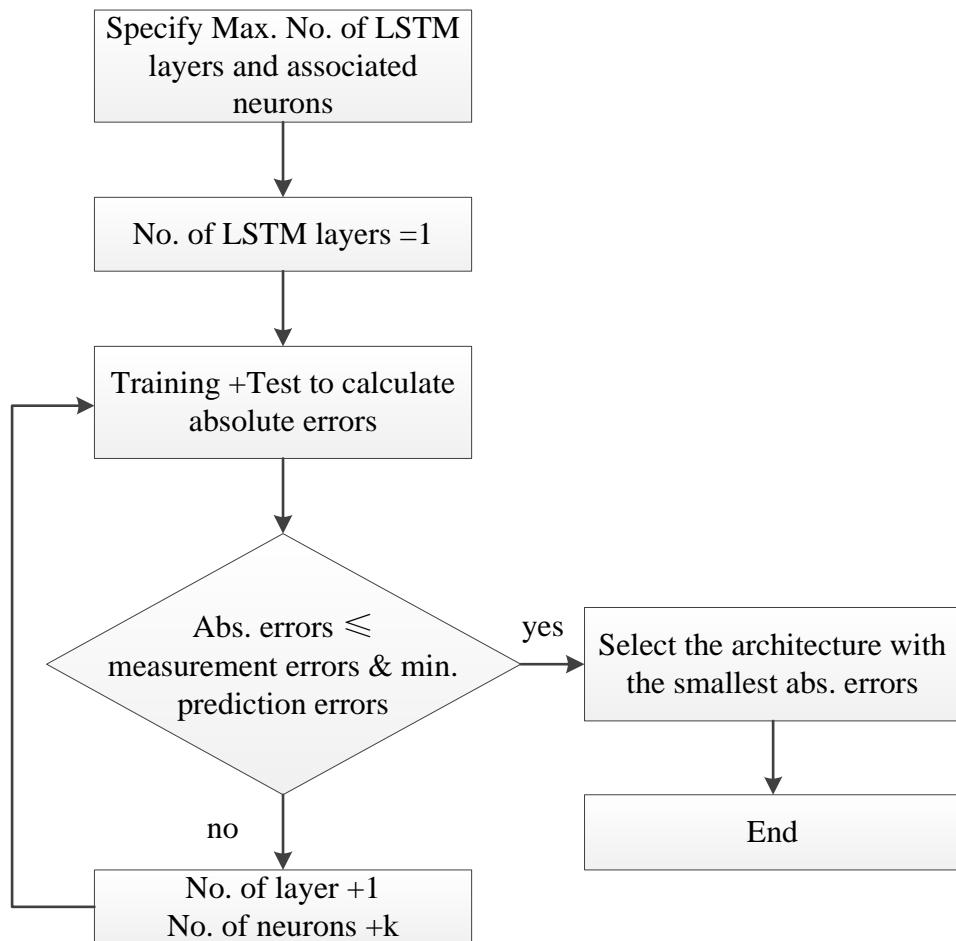


Figure 6. Flowchart of determining the network architecture.

2.3.3. Evaluation of the Optimal LSTM Network

An evaluation of the network was conducted in training set and testing set respectively. Overfitting problem happened when there was an extremely high accuracy in the training set but a low accuracy in the testing set happens. Oppositely, underfitting problem happened when both accuracies were low, but the testing one was higher than the training one [45,46]. Their accuracies were assessed to guarantee the model at the sweet spot between underfitting and overfitting, which ensures a good generalization performance.

Following statistical criteria was used to evaluate the accuracy of LSTM network [47]: coefficient of determination (R^2), mean square error (MSE), root mean square error ($RMSE$), mean absolute error (MAE) and standard deviation of the errors (StD of errors). These evaluation indicators were defined mathematically as follows:

$$\begin{aligned}
 R^2 &= 1 - \frac{\sum_{i=1}^m (y_{i(actual)} - \hat{y}_{i(predicted)})^2}{\sum_{i=1}^m (y_{i(actual)} - \bar{y}_{i(actual)})^2} \\
 MSE &= \frac{1}{m} \sum_{i=1}^m (y_{i(actual)} - y_{i(predicted)})^2 \\
 RMSE &= \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{i(actual)} - y_{i(predicted)})^2} \\
 MAE &= \frac{1}{m} \sum_{i=1}^m |y_{i(actual)} - y_{i(predicted)}| \\
 StD \text{ of errors} &= \sqrt{\frac{1}{(m-1)} \sum_{i=1}^m (E_{i(predicted)} - \bar{E}_{i(predicted)})^2}
 \end{aligned}$$

where $y_{(actual)}$ and $y_{(predicted)}$ are actual value and predicted value, m is the number of data, and E is the error between actual value and predicted value.

2.3.4. Practical Application of Real-Time Conformity Monitoring

The evaluation of the ANN model demonstrates the accuracy of micro-environmental prediction for the empty enclosure, which represents the worst-case scenario for hygrothermal buffering compared to enclosures containing paper-based collections. However, assessing the performance of the ANN model in practical application is essential. This part of the study aims to assess the robustness of real-time prediction for monitoring micro-environmental conformity in practical scenario where enclosures with the paper-based collections.

Validation Experiment

The trial was carried out in a test room, 5m long, 3m wide, and 3m tall (Figure 7). Inside the room, the air distribution was created to match that of the library's air-conditioning configuration, with air supplied from the top and returned from bottom, setting the velocity of supply air around 0.2m/s to prevent dust buoyancy. The macro-environment was controlled to fall within the range of 16~32°C and 30%~80%RH, with an accuracy of approximate $\pm 2^\circ\text{C}$ and $\pm 10\%$ RH. The operational points were set at 20°C and 55%RH for this trial.

The trial was carried out in the controlled chamber of 5m long, 3m wide and 3m tall (Figure 7). Inside this room, the conditioned air went into the room space through a void above a grid ceiling then returned at low level through three vents 0.1m above the floor. The desk panel was meant to separate two airflow zones. The upper zone was inside the main stream whilst the lower one experienced more turbulence. These configuration efforts were to replicate the air distribution of the real scenario for the validation purposes.

Two enclosures were strategically placed the two zones to account for temperature and humidity fluctuations across varying spatial gradients inside the space. They were filled with 70% full newspapers to represent the most common use of these enclosures. Additionally, two temperature and humidity loggers (loggers 2 & 4 in Figure 12) were placed inside these enclosures to record the micro-environment conditions. Another two loggers (loggers 1 & 3 in Figure 12) were positioned near these enclosures to monitor the macro-environment. Their measurement accuracies are $\pm 1^\circ\text{C}$ (within the range of 5-60°C) and $\pm 3\%$ RH (within the range of 20-80%RH). The sampling interval for all loggers was set at 15 minutes, consistent with the interval used in the BMS (Building Management System).

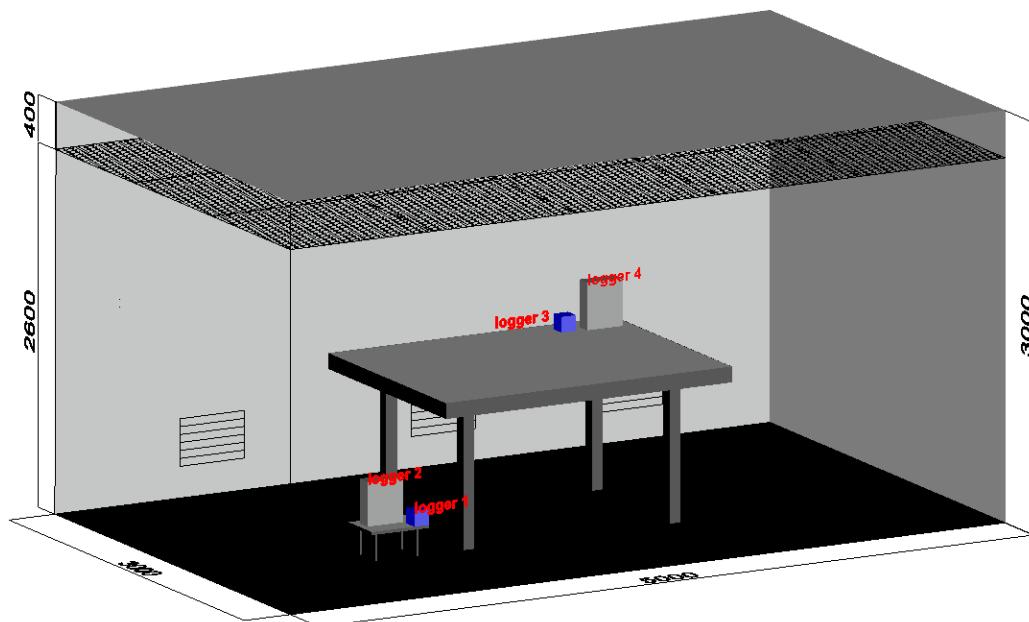


Figure 7. Diagram of test room (5m×3m×3m).

Parallel Prediction

To predict the micro-environment within various enclosures positioned differently, we replicated the trained ANN for each of those sensors located in the macro-environment. The trained ANN models were to act as parallel predictors in a BMS (Figure 8). The ANNs were fed with the latest 24h macro-environmental data, consisting of 96 data points with 15min sampling interval from the on-site monitoring, to predict the micro-environmental condition.

The maximum fluctuations and associated bands in the predicted results were calculated. Simultaneously, a real-time measurement of micro-environment was conducted to assess the prediction in the trial test. Analysing environmental data helps identify risks specific to the collections with various materials, but this inherent diversity makes this process complex [48]. Therefore, the ANN prediction was based on the worst-case scenario or least buffering effect, an empty box with no collections inside to further moderate the micro-environmental fluctuations. This resulted in a larger predicted fluctuations and bands compared to real-time measurements, which understandably led to deviations between the predicted and measured data. These deviations are defined as follows:

- 1) 24h Fluctuation_{*prediction*} – 24h Fluctuation_{*measurement*}
- 2) T/RH band_{*measurement*} ∈ T/RH band_{*prediction*} (the measurement bands are included in the prediction control ones).

These deviations provide a critical region between safe and dangerous conditions for preventive conservation. A positive value indicates that the actual micro-environment is more stable and safer than the predicted condition. Conversely, a negative value indicates that there is no critical region, and the actual conditions fall outside the scenarios captured by the trained ANN predictor, which triggers a built-in alarm. In such conditions, re-training the ANN with new measurement data is necessary.

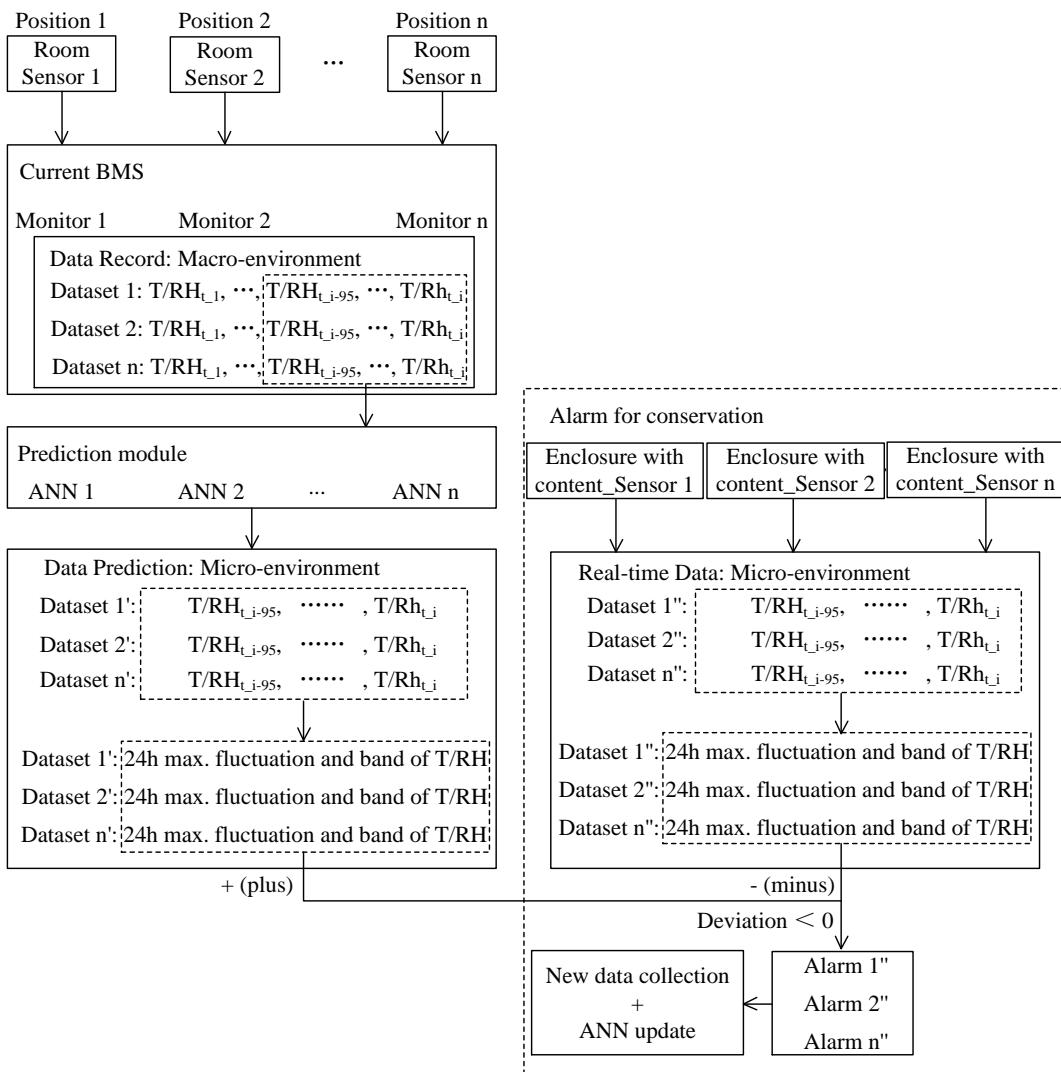


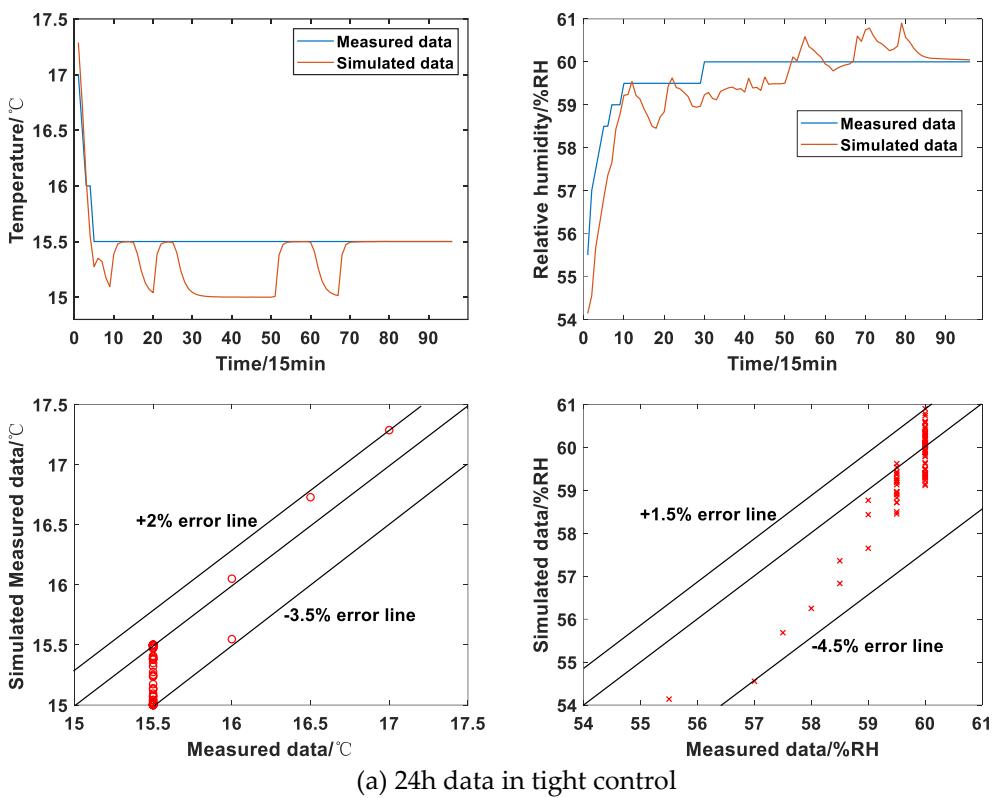
Figure 8. Framework of parallel prediction (“n” predictors for “n” sensors).

The test was run continuously for four days to assess the robustness of these ANN predictors worked in real-time. If any alarm triggered during the test, that ANN model should be retrained with new data to improve its robustness. Conversely, if no alarm was triggered, the robustness could be considered acceptable.

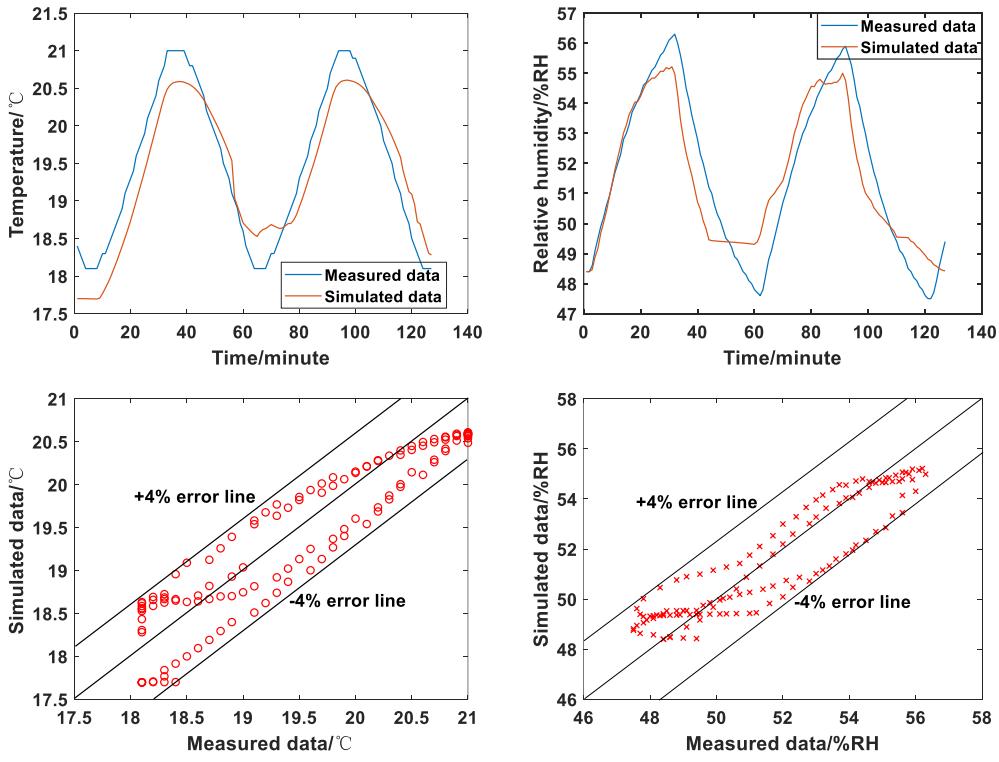
3. Results and Discussion

3.1. Comparison of Measured and Simulated Data in Heat and Mass Simulation

Comparison between the measured data and the simulated data for micro-environment temperature and RH shows relative errors within $\pm 4\%$ for temperature and $\pm 4.5\%$ for RH, and absolute errors below 0.7°C for temperature and 2.5% for RH (Figure 9). In details, the accuracy of the simulation is demonstrated through a series of comparative figures. The first figure a) illustrates the data for tight control condition, where the errors are kept below 3.5% for temperature and 4.5% for RH. The subsequent figures present the conditions under more relaxed control. The second figure b) depicts a scenario with a 5°C fluctuation over a 60-minute period at a fixed 50% RH, where the relative errors are under 4% for both temperature and RH. The third figure c) represents a 10% RH fluctuation over a 60-minute period at a fixed 20°C , with errors below 3% for temperature and 1% for RH. These results confirm the high accuracy of the simulation, validating their effectiveness in replicating real-world conditions.



(a) 24h data in tight control

(b) $\uparrow\downarrow 5^{\circ}\text{C}$ fluctuation during 60min period at fixed $\sim 50\%$ RH

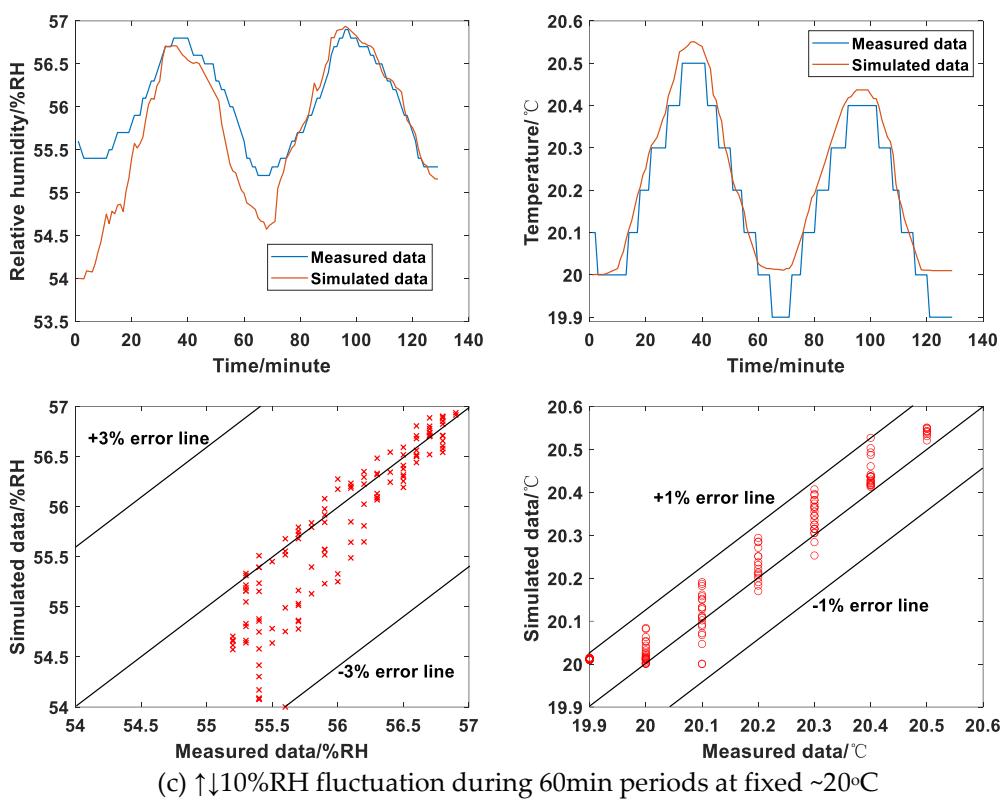


Figure 9. Comparison between measured data and simulated data.

To assess the agreement between measured and simulated data, Table 3 shows the KGE values in these three validation conditions. All values exceed the baseline ($\text{KGE} \geq 0.5$), which indicates an acceptable agreement in simulating the heat, air and moisture transfer between the micro-environment and macro-environment. The model effectively reproduces the variation trends and fluctuation periods in the measurement conditions for subsequent study about determination of acceptable macro-environment.

However, the results show relatively high values in the last two measurement conditions with the relaxing control, compared to the first condition with tight control. This suggests that the model exhibits greater agreement in the relaxing control than in the tight control.

Table 3. KGE values for micro-environmental temperature and RH.

	Tight-control	$\uparrow\downarrow 5^\circ\text{C} @ 50\%$ RH	$\uparrow\downarrow 10\%$ RH @ 20°C
KGE for temp.	0.51	0.84	0.97
KGE for RH	0.58	0.77	0.63

3.2. Acceptable Macro-Environment

The upper limit of acceptable macro-environmental conditions was determined using the trial-and-error method within the searching loop. The macro-environment can be relaxed to an upper limit of 33% to 65% RH with a $\pm 16\%$ RH fluctuation over 24 hours while the current requirement is 40% to 60% RH with a $\pm 10\%$ RH fluctuation.

Correspondingly, the micro-environment RH can be maintained within the range of 43% to 57.3%RH, with 24-hour fluctuation of $\leq \pm 9.1\%$ RH, as showed in Figure 10. These results illustrate that the buffering capacity of enclosure ensures a stable micro-environment with low RH fluctuation amplitude. This stability opens up a potential to relax current tight control in the macro-environment without any detrimental effect for the collections.

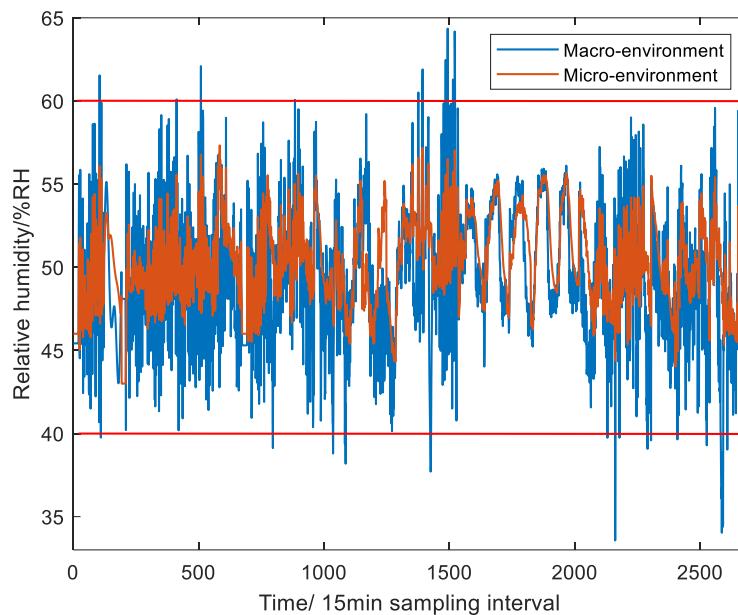


Figure 10. Acceptable macro and micro environmental RH in the most relaxing condition.

3.3. Real-Time Prediction of Micro-Environment

For data acquisition, four sets of data that were generated to represent different levels of macro-environmental control accuracy from $\pm 10\%$ RH 24h fluctuation to the most relaxing condition. They were prepared to illustrate the four levels of relaxing control within the acceptable macro-environment range, as showed in Figure 11. From the perspective of data expression, these datasets and the on-site dataset can reflect characteristics about the interaction of heat and mass transfer between macro and micro environments within the acceptable range.

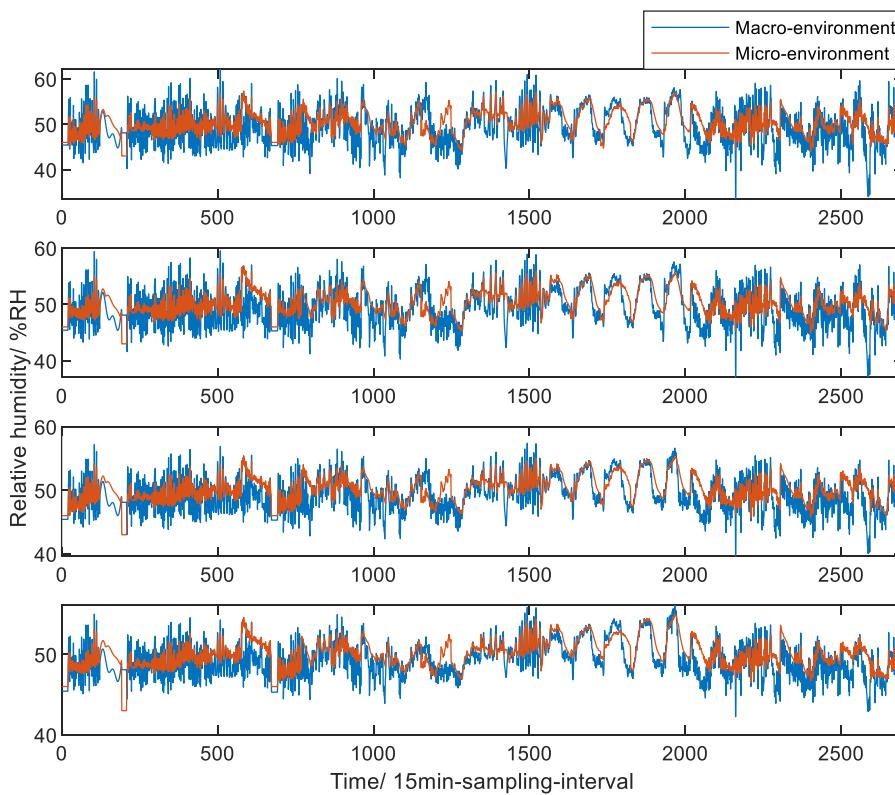


Figure 11. The macro and micro environments in four levels of control accuracy.

For searching the optimal architecture, multiple LSTM networks were trained through a gradual increase in the number of hidden layers and neurons. The resulting RMSE and absolute errors are detailed in Table 4. The RMSE values range from 0.037 to 0.188 for temperature and 0.467 to 1.73 for RH. Correspondingly, absolute errors vary from $|-0.13|$ to 0.64°C for temperature and $|-8.2\%|$ to 6.3% for RH.

The architecture with 2 layers demonstrates relatively high accuracy in prediction while a single-layer configuration shows relatively low accuracy in RH prediction. As the number of layers increases from 1 to 4 in the networks with an equal number of neurons per layer, accuracy typically improves. However, as the number of neurons increases in the networks with an equal number of layers, beyond a certain threshold where the number of neurons exceeds a maximum, further increasing the number of neurons per layer can lead to a decline in accuracy. Both a relatively shallow LSTM network with many neurons and a deeper network with fewer neurons demonstrate strong performance.

In detail, for the 2-layer networks, the configuration with 200 neurons in each layer stands out as the most accurate. Comparing its absolute errors with measurement errors ($\pm 0.5^{\circ}\text{C}$ and $\pm 3\%$ RH), it is noted that the absolute errors ranging from -0.13 to 0.15°C and -2.2% to 2.2% RH are smaller than the measurement errors. This suggests that the network can predict micro-environmental temperature and RH as accurately as on-site measurements.

Based on these findings, we conclude that the optimal LSTM network architecture consists of 2 layers, with each layer comprising 200 neurons.

Table 4. RMSE and absolute errors in the test set.

	50-neuron/layer	100-neuron/layer	150-neuron/layer	200-neuron/layer	250-neuron/layer
1 layer	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE value change</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">Absolute errors</div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: center;"> <p>0.188</p> <p>+1°C -1°C</p> <p>+0.44 -0.26</p> <p>+3%RH -3%RH</p> <p>1.73 -7.8</p> </div> <div style="text-align: center;"> <p>1.52</p> <p>+1°C -1°C</p> <p>+0.43 -0.24</p> <p>+3%RH -3%RH</p> <p>5.6 -6.9</p> </div> <div style="text-align: center;"> <p>MIN.</p> <p>0.131</p> <p>+1°C -1°C</p> <p>+0.38 -0.38</p> <p>+3%RH -3%RH</p> <p>4.6 -4.7</p> </div> <div style="text-align: center;"> <p>0.165</p> <p>+1°C -1°C</p> <p>+0.46 -0.2</p> <p>+3%RH -3%RH</p> <p>5.4 -8.2</p> </div> </div>				
2 layers	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE value change</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">Absolute errors</div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: center;"> <p>0.117</p> <p>+1°C -1°C</p> <p>+0.53 -0.15</p> <p>+3%RH -3%RH</p> <p>5.4 -6.9</p> </div> <div style="text-align: center;"> <p>1.54</p> <p>+1°C -1°C</p> <p>+0.28 -0.21</p> <p>+3%RH -3%RH</p> <p>5.4 -6.9</p> </div> <div style="text-align: center;"> <p>0.088</p> <p>+1°C -1°C</p> <p>+0.28 -0.21</p> <p>+3%RH -3%RH</p> <p>5.6 -5.7</p> </div> <div style="text-align: center;"> <p>0.058</p> <p>+1°C -1°C</p> <p>+0.21 -0.22</p> <p>+3%RH -3%RH</p> <p>3.5 -3.7</p> </div> <div style="text-align: center;"> <p>0.037</p> <p>+1°C -1°C</p> <p>+0.15 -0.13</p> <p>+3%RH -3%RH</p> <p>0.467 -2.2</p> </div> <div style="text-align: center;"> <p>0.071</p> <p>+1°C -1°C</p> <p>+0.21 -0.28</p> <p>+3%RH -3%RH</p> <p>5.0 -4.9</p> </div> </div>				
3 layers	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE value change</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">Absolute errors</div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: center;"> <p>0.081</p> <p>+1°C -1°C</p> <p>+0.41 -0.14</p> <p>+3%RH -3%RH</p> <p>5.8 -6.0</p> </div> <div style="text-align: center;"> <p>1.421</p> <p>+1°C -1°C</p> <p>+0.36 -0.28</p> <p>+3%RH -3%RH</p> <p>5.8 -6.0</p> </div> <div style="text-align: center;"> <p>MIN.</p> <p>0.071</p> <p>+1°C -1°C</p> <p>+0.36 -0.28</p> <p>+3%RH -3%RH</p> <p>5.2 -6.7</p> </div> <div style="text-align: center;"> <p>0.141</p> <p>+1°C -1°C</p> <p>+0.42 -0.31</p> <p>+3%RH -3%RH</p> <p>5.9 -6.2</p> </div> </div>				
4 layers	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE value change</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">Absolute errors</div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: center;"> <p>0.121</p> <p>+1°C -1°C</p> <p>+0.30 -0.51</p> <p>+3%RH -3%RH</p> <p>3.2 -3.1</p> </div> <div style="text-align: center;"> <p>0.867</p> <p>+1°C -1°C</p> <p>+0.42 -0.20</p> <p>+3%RH -3%RH</p> <p>3.2 -6.2</p> </div> </div>				
5 layers	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE value change</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">RMSE</div> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">Absolute errors</div> </div> <div style="display: flex; justify-content: space-between;"> <div style="text-align: center;"> <p>0.147</p> <p>+1°C -1°C</p> <p>+0.43 -0.44</p> <p>+3%RH -3%RH</p> <p>4.1 -4.1</p> </div> <div style="text-align: center;"> <p>1.123</p> <p>+1°C -1°C</p> <p>+0.43 -0.44</p> <p>+3%RH -3%RH</p> <p>4.1 -4.1</p> </div> </div>				

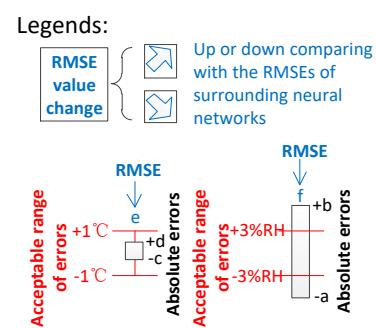
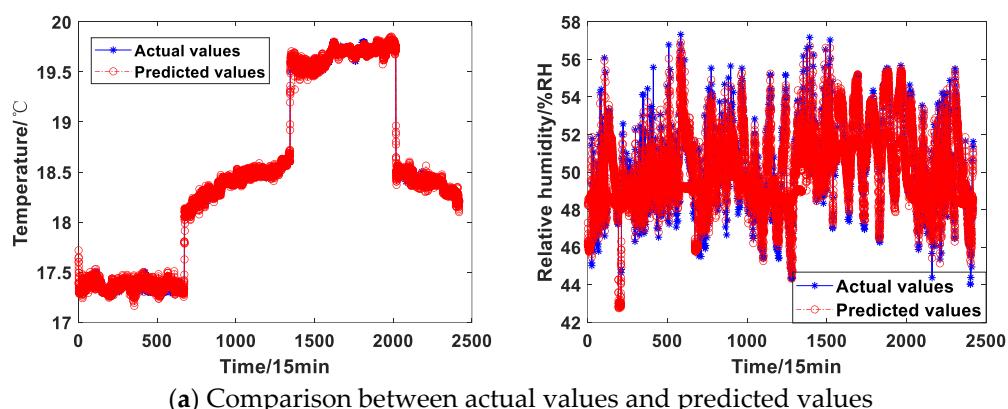


Table 5 shows the performance results of the optimal LSTM network. All coefficients of determination are close to 1 while the values of MSE, RMSE, MAE and StD are near 0. In addition,

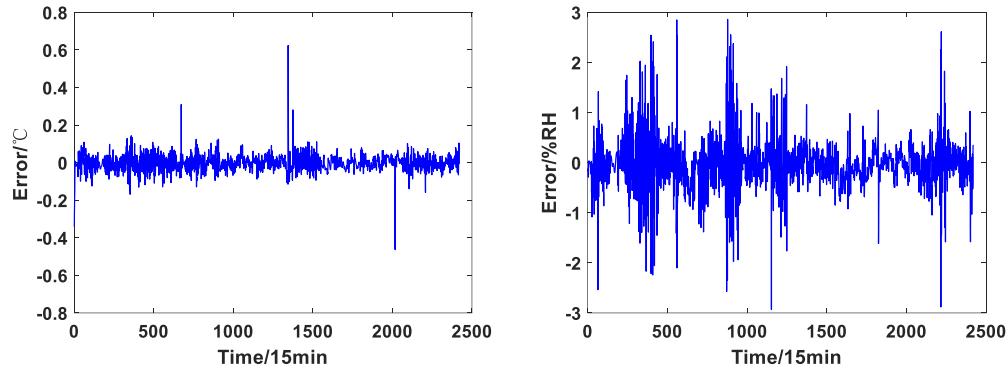
Figures 12 and 13 illustrate: 1) the time-series prediction adeptly capture the fluctuation trends with accurate periods in both training and testing sets; 2) the absolute errors predominantly remain below $\pm 0.2^{\circ}\text{C}$ and $\pm 3\%$ RH in the training set, and $\pm 0.1^{\circ}\text{C}$ and $\pm 2\%$ RH in testing set; 3) the error histograms present a distribution centred around zero with minimal deviation. These indicators indicate high accuracy in both training and testing processes. They affirm that this optimal network has robust generalization and accurate prediction performance without underfitting and overfitting problem.

Table 5. Performance results of the optimal LSTM network.

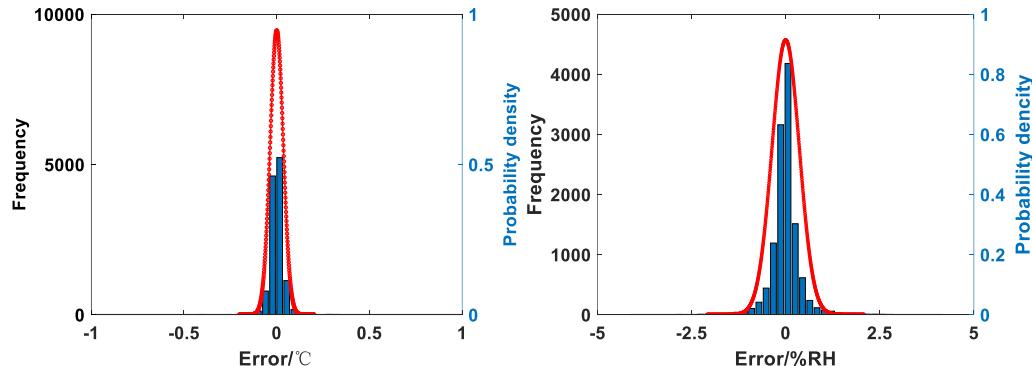
	R ²	RMSE	MSE	MAE	StD
Training data (temperature)	0.999	0.035	0.001	0.023	0.035
Training data (RH)	0.984	0.364	0.132	0.230	0.364
Testing data (temperature)	0.965	0.037	0.001	0.025	0.037
Testing data (RH)	0.963	0.468	0.219	0.313	0.468



(a) Comparison between actual values and predicted values



(b) Errors



(c) Error histogram

Figure 12. Evaluation of the optimal LSTM in the train set.

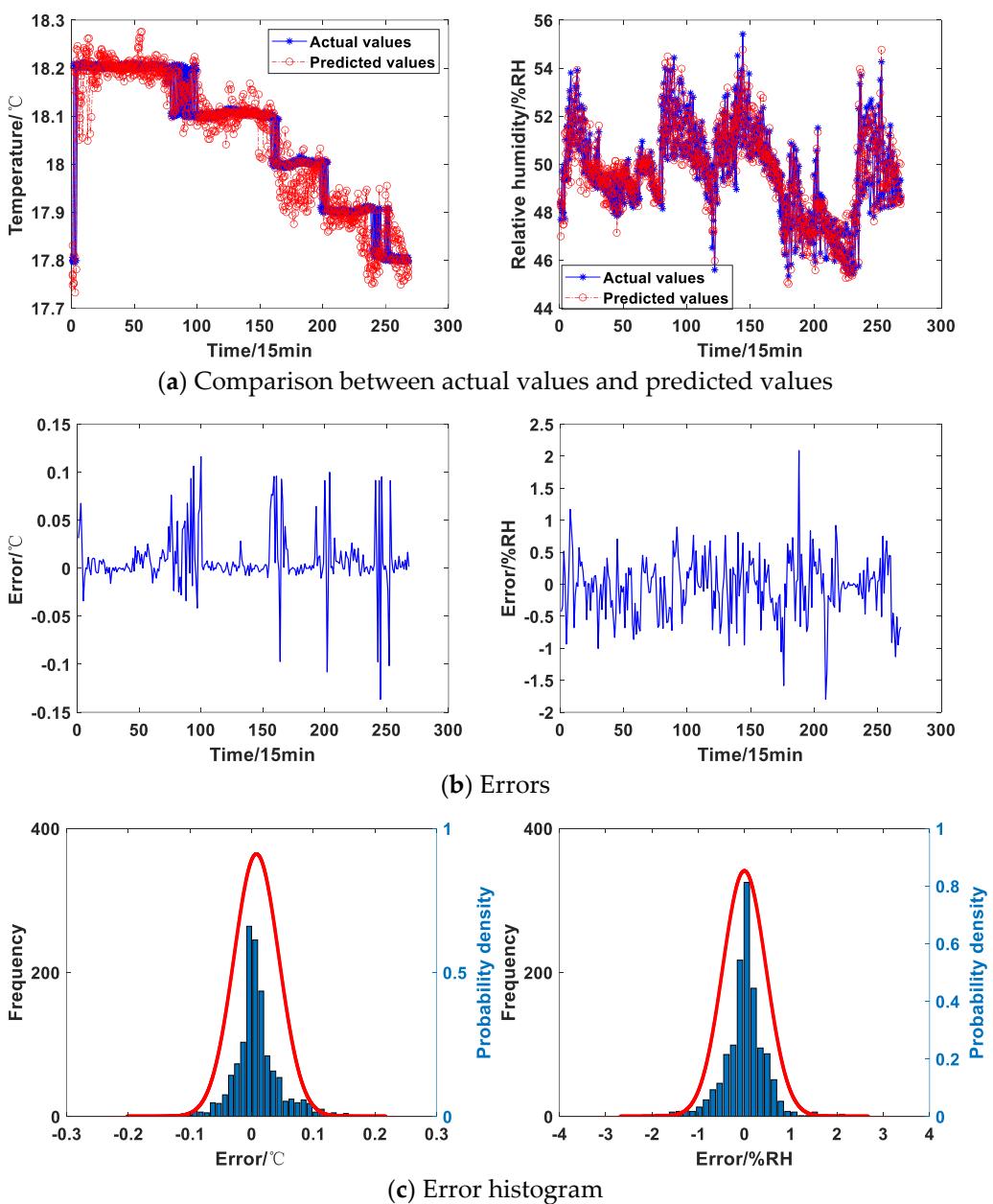


Figure 13. Evaluation of the optimal LSTM in the test set.

Additionally, the prediction time for daily data, comprising 96 data points, was less than 5 seconds by using a 1.6GHz CPU with 16GB RAM. Based on this trained LSTM network, any 24h time-sequence macro-environmental temperature and RH data can be input to predict the corresponding micro-environmental ones inside the storage enclosure.

The model evaluation above confirmed the accuracy of micro-environmental prediction under the worst-case scenario without any buffering effect from paper-based collections. The following results focus on assessing the model robustness in practical applications involving enclosures with collections inside.

Figure 14 shows data comparison between onsite measurement and ANN prediction. Notably, the micro-environment in return air zone exhibits better stability compared to that in the supply air zone. Associated absolute errors are $\pm 0.6^{\circ}\text{C}$ & $\pm 1.6\%\text{RH}$ for the return air zone and $\pm 0.9^{\circ}\text{C}$ & $\pm 3\%\text{RH}$ for the supply air zone. These prediction accuracies are smaller than the onsite monitoring accuracy of $\pm 1^{\circ}\text{C}$ and $\pm 3\%\text{RH}$, which is deemed acceptable for the real-time micro-environment prediction.

In addition, the temperature and humidity fluctuations in return air zone range from 0.75 to 1°C and 1 to 11%RH, while the supply air zone has the fluctuations of 2.5 to 3°C and 4 to 13%RH. Despite

the larger fluctuations in the supply air zone, deviations of both zones are negative value with the critical region of approximate 0.3°C and $1\%\text{RH}$. It means that the measurement temperature and RH bands are encompassed within the prediction ones for any 24h period, thus triggering no alarms.

This real-time prediction provides a feasible way to monitor the micro-environment, and to map the hygrothermal condition over multitude enclosures in the macro environment, the storage room. This would easily allow us to test relaxing the macro-environmental control to reduce operational energy consumption of the precision air-conditioning system without risk of damaging collection by doing it in the reality.

Consequently, this ANN prediction is robust in conformity monitoring of the micro-environment and protecting the collections stored inside the enclosures. Furthermore, this prediction contributes to implementation of the relaxing control in macro-environment, thereby promoting sustainable preventive conservation practices.

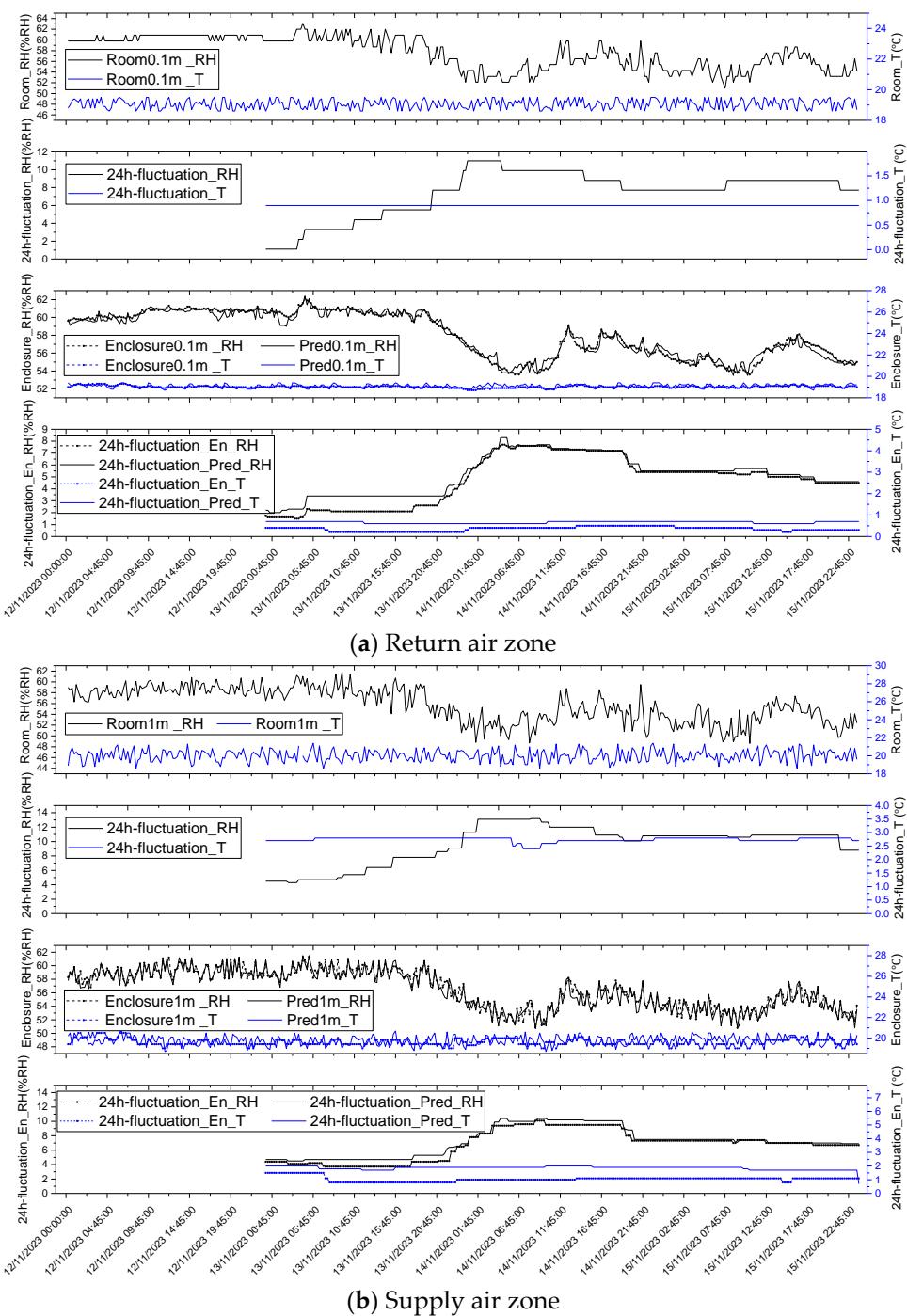


Figure 14. Data comparison between onsite measurement and ANN prediction.

4. Conclusions

The storage enclosure offers a buffering capacity that moderates micro-environmental temperature and RH fluctuations, enabling energy savings by relaxing macro-environmental control in the room. Nevertheless, there is a practical barrier in implementing physical sensors for each enclosure to achieve a real-time conformity monitoring of the micro-environment to ensure compliance of preventive conservation specifications. To address this limitation, this study developed a coupled heat and mass transfer model. This model served to establish acceptable macro-environmental conditions while ensuring compliance of the micro-environment with preventive conservation specifications. Additionally, the data generated from the heat and mass simulation were utilized to train an ANN model. This ANN model was successfully deployed in a real-world application, demonstrating promising results for real-time conformity monitoring of the micro-environment. The key conclusions are outlined as follows.

1) The coupled heat and mass model effectively captured the hygrothermal interaction between macro and micro environments with high accuracy and acceptable agreement. It can reproduce the variation trends and fluctuation periods in the measurement conditions to determine acceptable macro-environment.

2) The acceptable macro-environment can be relaxed from 40% to 60% RH with a $\pm 10\%$ RH 24-h fluctuation to 3% to 65% RH with a $\pm 16\%$ RH fluctuation, while ensuring the compliance of micro-environment with the specifications of preventive conservation.

3) The ANN-based prediction can achieve highly accuracy 24h time-sequence prediction of micro-environment within 5 seconds. This prediction demonstrates satisfactory robustness in real-time conformity monitoring, further facilitating implementation of the relaxing control strategy for sustainable preventive conservation practices.

These efforts leverage the advantage of ANN-based approach to overcome the limitation about real-time conformity monitoring of micro-environment without disrupting on-site operation of the air conditioning system in the NLS. In a practical application, using real-time prediction for the micro-environment temperature and RH eliminates the necessity for monitoring sensors inside each enclosure and offers an easy way to map the micro environmental conditions inside multitude of enclosures in the large storage space. This particular predictive capacity could enable monitoring the micro-environment within the enclosures against the desirable condition under various control accuracy in the storage room, the macro-environment, consequently testing the control levels that saves energy in the service system and safe to the collections inside the enclosures without detrimental effect of the collections stored inside the enclosures.

While this study provides valuable insights, it is based on a limited dataset for ANN training. Future work should focus on continuous data training, updating the current network with new data to enhance prediction accuracy and ensure the safeguarding of collection care. This modelling approach offers a promising solution for monitoring micro-environmental conformity and implementing a relaxed control strategy in the macro-environment without compromising the integrity of the collections stored inside the enclosures.

Funding: This work is a part of three-year PhD study, funded by the Energy Technology Partnership, Scotland (ETP 173 - 2019).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Muñoz-Viñas, S. (2012). Contemporary theory of conservation. Routledge.
2. Wang, Y., Wang, Y., Chen, Q., Lei, Y., & Tang, K. (2025). Identification, deterioration, and protection of organic cultural heritages from a modern perspective. *Heritage Science*, 13(1), 1-19.

3. Elnaggar, A., Said, M., Kraševc, I., Said, A., Grau-Bove, J., & Moubarak, H. (2024). Risk analysis for preventive conservation of heritage collections in Mediterranean museums: case study of the museum of fine arts in Alexandria (Egypt). *Heritage Science*, 12(1), 1-17.
4. Ganguly, S., Wang, F., & Browne, M. (2019). Comparative methods to assess renovation impact on indoor hygrothermal quality in a historical art gallery. *Indoor and Built Environment*, 28(4), 492-505.
5. Wang, F., & Ganguly, S. (2015). Effects of Renovation Solutions in a Historical Building of Cultural Significance.
6. Fernandez-Cortes, A., Palacio-Perez, E., Martin-Pozas, T., Cuevva, S., Ontañon, R., Lario, J., & Sanchez-Moral, S. (2025). Defining the Optimal Ranges of Tourist Visits in UNESCO World Heritage Caves with Rock Art: The Case of El Castillo and Covalanas (Cantabria, Spain).
7. Harrestrup, Maria, and Svend Svendsen. "Internal insulation applied in heritage multi-storey buildings with wooden beams embedded in solid masonry brick façades." *Building and environment* 99 (2016): 59-72.
8. Shiner, J. (2007). Trends in microclimate control of museum display cases. *Museum Microclimates*, 267-275.
9. Cao, S., Ge, W., Yang, Y., Huang, Q., & Wang, X. (2022). High strength, flexible, and conductive graphene/polypropylene fiber paper fabricated via papermaking process. *Advanced Composites and Hybrid Materials*, 1-9.
10. Zhao, C., Mark, L. H., Chang, E., Chu, R. K., Lee, P. C., & Park, C. B. (2020). Highly expanded, highly insulating polypropylene/polybutylene-terephthalate composite foams manufactured by nano-fibrillation technology. *Materials & Design*, 188, 108450.
11. Jose, J., Thomas, V., Raj, A., John, J., Mathew, R. M., Vinod, V., ... & Mujeeb, A. (2020). Eco-friendly thermal insulation material from cellulose nanofibre. *Journal of Applied Polymer Science*, 137(2), 48272.
12. Gane, P. A., & Ridgway, C. J. (2009). Moisture pickup in calcium carbonate coating structures: role of surface and pore structure geometry. *Nordic Pulp & Paper Research Journal*, 24(3), 298-308.
13. Derluyn, H., Janssen, H., Diepens, J., Derome, D., & Carmeliet, J. (2007). Hygroscopic behavior of paper and books. *Journal of Building Physics*, 31(1), 9-34.
14. Han, B., Li, X., Wang, F., Bon, J., & Symonds, I. (2024). The buffering effect of a paper-based storage enclosure made from functional materials for preventive conservation. *Indoor and Built Environment*, 33(1), 167-182.
15. Perino, M. (2018). Air tightness and RH control in museum showcases: Concepts and testing procedures. *Journal of Cultural Heritage*, 34, 277-290.
16. van Schijndel, A. W. M. (2007). Integrated heat air and moisture modeling and simulation.
17. Simonson, C. J., Salaonvaara, M., & Ojanen, T. (2004). Heat and mass transfer between indoor air and a permeable and hygroscopic building envelope: part II-verification and numerical studies. *Journal of Thermal Envelope and Building Science*, 28(2), 161-185.
18. Dong, W., Chen, Y., Bao, Y., & Fang, A. (2020). A validation of dynamic hygrothermal model with coupled heat and moisture transfer in porous building materials and envelopes. *Journal of Building Engineering*, 32, 101484.
19. Singh, S., & Abbassi, H. (2018). 1D/3D transient HVAC thermal modeling of an off-highway machinery cabin using CFD-ANN hybrid method. *Applied Thermal Engineering*, 135, 406-417.
20. Darvishvand, L., Safari, V., Kamkari, B., Alamshenas, M., & Afrand, M. (2022). Machine learning-based prediction of transient latent heat thermal storage in finned enclosures using group method of data handling approach: a numerical simulation. *Engineering Analysis with Boundary Elements*, 143, 61-77.
21. Ganguly, S., Ahmed, A., & Wang, F. (2020). Optimised building energy and indoor microclimatic predictions using knowledge-based system identification in a historical art gallery. *Neural Computing and Applications*, 32(8), 3349-3366.
22. Steeman, H. J., Van Belleghem, M., Janssens, A., & De Paepe, M. (2009). Coupled simulation of heat and moisture transport in air and porous materials for the assessment of moisture related damage. *Building and Environment*, 44(10), 2176-2184.
23. CEN, E. (2007). 15026: Hygrothermal Performance of Building Components and Building Elements-Assessment of Moisture Transfer by Numerical Simulation. CEN, Brussels, Belgium.

24. Tariku, F., Kumaran, K., & Fazio, P. (2010). Transient model for coupled heat, air and moisture transfer through multilayered porous media. *International journal of heat and mass transfer*, 53(15-16), 3035-3044.
25. Li, L., Zhang, C., Xu, X., Yu, J., Wang, F., Gang, W., & Wang, J. (2022, July). Simulation study of a dual-cavity window with gravity-driven cooling mechanism. In *Building Simulation* (pp. 1-14). Tsinghua University Press.
26. Liu, D. (2020). A rational performance criterion for hydrological model. *Journal of Hydrology*, 590, 125488.
27. Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23(10), 4323-4331.
28. Lee, K., Yoo, H., & Levermore, G. J. (2010). Generation of typical weather data using the ISO Test Reference Year (TRY) method for major cities of South Korea. *Building and Environment*, 45(4), 956-963.
29. International Organization for Standardization (2005). ISO 15927-4: Hourly data for assessing the annual energy use for heating and cooling. ISO, Geneva, Switzerland.
30. Grass, A. M., & Gibbs, F. A. (1938). A Fourier transform of the electroencephalogram. *Journal of neurophysiology*, 1(6), 521-526.
31. Sanei, S., & Chambers, J. A. (2013). *EEG signal processing*. John Wiley & Sons.
32. Sawant, H. K., & Jalali, Z. (2010). Detection and classification of EEG waves. *Oriental Journal of Computer Science and Technology*, 3(1), 207-213.
33. Nussbaumer, H. J., & Nussbaumer, H. J. (1982). The fast Fourier transform (pp. 80-111). Springer Berlin Heidelberg.
34. British Standards Institution. (2012). Specification for Managing Environmental Conditions for Cultural Collections: PAS 198: 2012. British Standards Limited.
35. Qiao, D., Li, P., Ma, G., Qi, X., Yan, J., Ning, D., & Li, B. (2021). Realtime prediction of dynamic mooring lines responses with LSTM neural network model. *Ocean Engineering*, 219, 108368.
36. Granata, F., Di Nunno, F., & de Marinis, G. (2022). Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study. *Journal of Hydrology*, 613, 128431.
37. Arslan, N., & Sekertekin, A. (2019). Application of Long Short-Term Memory neural network model for the reconstruction of MODIS Land Surface Temperature images. *Journal of Atmospheric and Solar-Terrestrial Physics*, 194, 105100.
38. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
39. Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, 31(7), 1235-1270.
40. Gal, Y., & Ghahramani, Z. (2016). A theoretically grounded application of dropout in recurrent neural networks. *Advances in neural information processing systems*, 29.
41. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
42. Ma, W., & Lu, J. (2017). An equivalence of fully connected layer and convolutional layer. *arXiv preprint arXiv:1712.01252*.
43. Muppidi, S., & PG, O. P. (2023). Dragonfly political optimizer algorithm-based rider deep long short-term memory for soil moisture and heat level prediction in iot. *The Computer Journal*, 66(6), 1350-1365.
44. Song, W., Gao, C., Zhao, Y., & Zhao, Y. (2020). A time series data filling method based on LSTM—Taking the stem moisture as an example. *Sensors*, 20(18), 5045.
45. Jabbar, H., and Rafiqul Zaman Khan. "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)." *Computer Science, Communication and Instrumentation Devices* 70 (2015): 163-172.
46. Bilbao, I., & Bilbao, J. (2017, December). Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks. In *2017 eighth international conference on intelligent computing and information systems (ICICIS)* (pp. 173-177). IEEE.
47. Althelaya, K. A., El-Alfy, E. S. M., & Mohammed, S. (2018, April). Evaluation of bidirectional LSTM for short-and long-term stock market prediction. In *2018 9th international conference on information and communication systems (ICICS)* (pp. 151-156). IEEE.

48. Perles, A., Fuster-López, L., & Bosco, E. (2024). Preventive conservation, predictive analysis and environmental monitoring. *Heritage Science*, 12(1), 11.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.