

PAPER • OPEN ACCESS

## Comparison of grey-box model and artificial neural network – prediction of surface condensation in residential space

To cite this article: E J Ju *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **609** 032016

View the [article online](#) for updates and enhancements.

You may also like

- [Modeling Protonic-Ceramic Fuel Cells with Porous Composite Electrodes in a Button-Cell Configuration](#)  
Huayang Zhu and Robert J. Kee
- [Mass comparisons using air buoyancy artefacts](#)  
A Picard and H Fang
- [Simultaneous Measurement of Air Temperature and Humidity Based on Sound Velocity and Attenuation Using Ultrasonic Probe](#)  
Takahiro Motegi, Koichi Mizutani and Naoto Wakatsuki

**PRIME**  
PACIFIC RIM MEETING  
ON ELECTROCHEMICAL  
AND SOLID STATE SCIENCE

HONOLULU, HI  
Oct 6–11, 2024

Abstract submission deadline:  
**April 12, 2024**

Learn more and submit!

**Joint Meeting of**  
The Electrochemical Society  
•  
The Electrochemical Society of Japan  
•  
Korea Electrochemical Society

# Comparison of grey-box model and artificial neural network – prediction of surface condensation in residential space

E J Ju<sup>1</sup>, J H Lee<sup>1</sup>, S H Park<sup>2</sup>, C S Park<sup>1</sup> and M S Yeo<sup>1,\*</sup>

<sup>1</sup> Department of Architecture and Architectural Engineering, Seoul National University, South Korea

<sup>2</sup> Data Science Lab, Korea Electric Power Corporation (KEPCO), South Korea

\* msyeo@snu.ac.kr

**Abstract** To apply real-time predictive control using automated devices for minimizing the risk of surface condensation in a residential space, the authors first developed a nodal network model that simulates the flow of moist air and the thermal behavior of a target area with the given boundary conditions of a space. The lumped model was enhanced using a parameter estimation technique based on the measured temperature, humidity, and schedule data. However, the humidity model prediction performance was still outside the valid range. A data-driven model was then developed using an artificial neural network (ANN) with the measured data that was formerly used to enhance the lumped model. Taking into consideration the possible uncertain characteristics of moist air, it was found that the data-driven model was a more suitable option for predicting the condensation as compared to the physics-based and grey-box models. With a stable range of errors between the simulation outputs and measured data, the ANN model could be useful for model predictive control.

## 1. Introduction

As buildings become airtight, the problem of controlling moisture has become a significant issue. Condensation is a typical example of bad control with respect to indoor moisture [1,2]. Residential buildings necessarily include wet areas such as a bathroom or a kitchen, where moisture is generated intermittently, thus making the adjacent spaces vulnerable to condensation. Condensation resulting from excessive moisture can result in poor indoor air quality, mold growth, or even structural damage [3]. As moisture transfer is a rapid process, and the consequent condensation phenomenon involves phase change [4], predictive control rather than corrective control is required to minimize the condensation risk [5]. Therefore, indoor moisture prediction based on simulation models should be performed prior to the establishment of control strategies [6,7].

Two different approaches exist for the simulation model: physics-based and data-driven. The physics-based model, also known as the forward approach, is based on the governing equations of building physics and provides a high generalization performance [7]. Computational fluid dynamics and the nodal network model are examples for the airflow simulations [8]. However, detailed information regarding the target space is required but unknown parameters can exist in a complex building system [9]. Moreover, the data-driven model describes the dynamics of a building with a relationship between the input and output data [10]. Even without making the additional effort to establish an assumption or perform a calibration step based on expertise, this model exhibits a high accuracy. The grey-box model is a combination of the physics-based and data-driven models, as its structure is based on the first principles and the unknown parameters are obtained from the measured data [11].

In this study, the authors developed prediction models comprising condensation determination factors using a nodal network and machine learning in model predictive control (MPC). Both these models predict the heat and mass change at a specific dry space in a residential building: a dressing room that is exposed to moisture transferred from a bathroom. It is assumed to have a lumped state in the



physics-based model to ensure a fast computation for real-time predictive control, and the parameter estimation technique was adopted with the use of measured data to improve the accuracy of this simplified model [13]. A machine learning model was developed using an artificial neural network (ANN) with the same data used for the grey-box model development. The data measurement, model development, and the comparison between these two different approaches are presented in the following section.

## 2. Analysis of surface condensation with field measurement

Surface condensation occurs when the temperature of a surface is lower than the dew point temperature calculated from the temperature and relative humidity. This means that there are two possible causes for surface condensation: excessive moisture or low surface temperature. Residential buildings thus belong to a high-risk group in terms of condensation as moisture is generated in their interiors [1]. To determine the actual occurrence patterns, the authors performed field measurements in winter. Rapid fluctuation was observed in the dew point temperature owing to the transferred moisture, and the requirement of a prediction model emerged for the complete prevention of surface condensation (Figure 1).

Data was collected from a testbed of apartment housing located in the Seoul metropolitan region, Uijeongbu City. This apartment was selected as the testbed because it has a type of floor plan that is significantly affected by interior moisture generation; the bathroom is in the interior zone, and the adjacent dressing room is facing north [2] as shown in Figure 2. The state variables were measured at 1-min intervals with at least two points in each area to clarify the causality of the data. To simulate the usage of a shower-booth faucet, a solenoid valve was installed in the bathroom. Furthermore, a pre-installed ventilation fan and bathroom door were set as control variables with actuators. Prior to the application of MPC, to determine the actual control effects with the change in control variables, various measurement scenarios were tested under different control logics. The measured parameters are listed in Table 1.

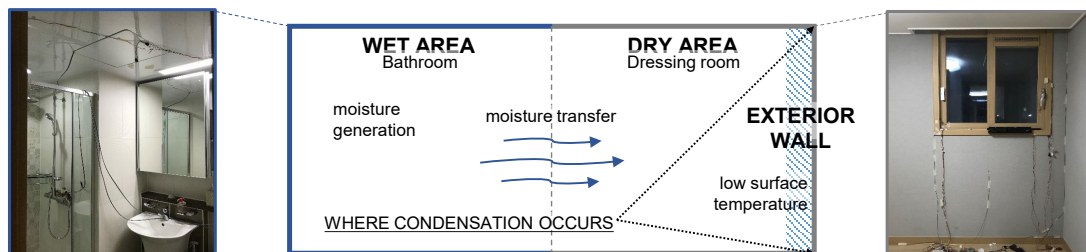


Figure 1. How surface condensation occurs in a residential space

Table 1. Measured item and location

Measured variables		Location	Description
State variables	Temperature(°C)	Dressing room	x1~2
		Bathroom	x3~4
		Intermediate zone	x5~6
	Relative humidity(%)	Dressing room	x7~8
		Bathroom	x9~10
		Intermediate zone	x11~12
	Surface temperature(°C)	North wall surface	x13~17
Openness(Open/Close)	Bathroom door/bedroom door	x18~19	
	Dressing room window	x20	
Control variables	Solenoid valve state(On/Off)	Bathroom shower booth	x21
	Swing door operator state(On/Off)	Bathroom door	x22
	Ventilation fan state(On/Off)	Bathroom	x23

### 3. Surface condensation prediction model development

#### 3.1. Physics-based and grey-box model

The nodal network model is based on the state space equation which explains the exchange between nodes. In this study, the heat balance and moisture balance equations are coupled for the calculation of the absolute humidity based on the temperature and relative humidity in the bathroom and dressing room. The basic structure of the airflow model depends on the moisture exchange between the spaces and infiltration from outdoor air, which is given by Equation 1. At the bathroom node, the airflow owing to the ventilation fan and the moisture generation were also modeled. Some terms that are considered in a typical calculation are ignored here, such as the moisture storage terms of walls, because vinyl wallpapers are used as wall finishing materials in this building, and its sorption effect is negligible. To speed up the computational time, the walls between spaces were set as space-averaged nodes as the aforementioned “lumped” states in the thermal calculation [12,13]. Consequently, the heat balance equation was expressed in terms of an overall heat transfer coefficient and airflow as per Equation 2 with an average U value of the walls. The surface temperature in the corner, as a separated node from the whole wall construction with part thermal capacity, was then calculated using the outdoor air temperature, and the calculation results of the heat balance equation according to the RC network are presented in Figure 3.

$$\rho V_z \frac{dW_z}{dt} = \sum \dot{m}_{gen} + \sum \dot{m}_i (W_i - W_z^t) + \dot{m}_{inf} (W_\infty - W_z^t) \quad (1)$$

$$\rho V_z \frac{dT_z}{dt} = \sum \dot{Q} + \sum U_i A_i (T_{i,wall} - T_z^t) + \sum \dot{m}_i C_p (T_i - T_z^t) + \dot{m}_{inf} C_p (T_\infty - T_z^t) \quad (2)$$

To compensate the degradation in the accuracy due to simplification, the parameter estimation technique was applied. It was found from data analysis that the lower corner surface temperature (x15) was the lowest, and the corresponding measured data for the spatial air (temperature x1, x4, and x5; relative humidity x7, x10, and x11) were sorted to identify uncertain parameters in order to improve the accuracy of the models. The outdoor climate data including the temperature, relative humidity, and barometric pressure were acquired from the online database of the Korea Meteorological Administration.

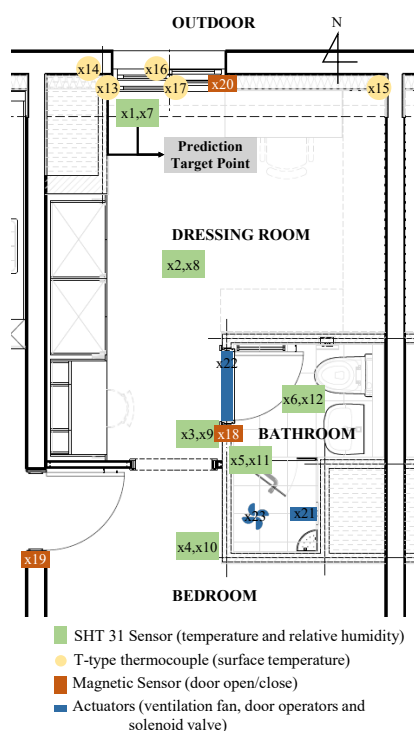
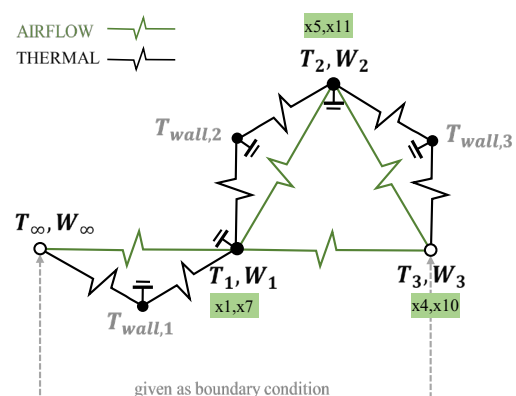
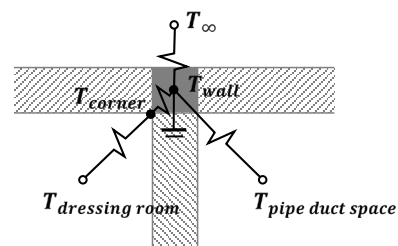


Figure 2. Floor plan of the testbed



(a) Heat and mass transfer in spatial air



(b) Corner surface thermal transmittance  
Figure 3. RC network diagram

The nodal network model was thus trained to minimize the mean-square error with the measured data as the grey-box model distinguished from the physics-based model. From the 20,000 data points obtained after the elimination of outliers, 15,000 points were used to train, and the other 5,000 points were used for the test. As a result, the performance of the model was slightly improved as shown in Table 2 and Figures 5 and 6. However, the relative humidity model still did not satisfy the ASHRAE guideline. This is because the re-evaporation effect after moisture generation and the transferred amount through the ventilation fan duct from neighboring apartment units could not be quantified using the model structure. Furthermore, not all the boundary conditions were known or measured; for example, exterior convective heat transfer coefficient changes consistently and temperature in pipe duct space is still unknown, and therefore the predicted values of the surface temperature have lost in tendency.

### 3.2. Data-driven model

Among various data-driven methods, ANN is broadly used for various purposes without being limited by materials. Furthermore, it exhibits a high accuracy in predicting nonlinear systems such as buildings and has a short computational time [6,7]. ANN learns from data in the process of minimizing errors between the actual data and its predicted values using back propagation technique, and as a result, finds the optimal weight of the neurons. The constraints can be expressed as follows, where  $y(w, x_j)$  and  $d_j$  represent the values predicted by the ANN and the measured data, respectively:

$$\operatorname{argmin} E(w) = \frac{1}{N} \sum_{j=1}^N (y(w, x_j) - d_j)^2 \quad (3)$$

In the ANN model, the prediction time horizon and target data must be predetermined at the training stage. In this study, the ANN was designed to predict the temperature, relative humidity, and surface temperature for 5 min of the future at the target point in Figure 2, i.e., the dressing room window (x1 and x7) and wall surface temperature (x15), using the state variables of the past 5 min including the states at the target point, and the future 5 min of the control variables that cause change. The time horizon was set as 5 min because it was the minimum control time step that could be used while considering a typical shower time. Data for the same period was used for the grey-box model. However, the difference was that the whole dataset (x1~x23) was used to train the ANN model because each data exhibited a strong positive correlation with at least one output variable. This is because of the good mixing of the airflow.

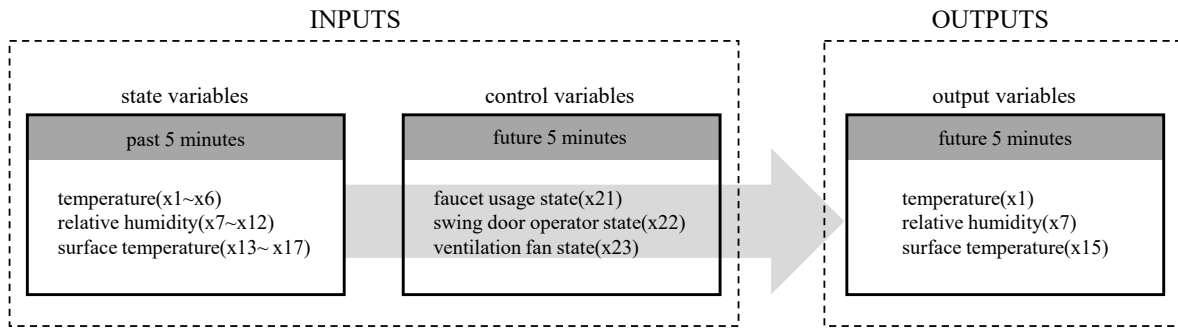
The structure of the model consists of two hidden layers. The number of nodes of the hidden layers and hyper parameters were determined using a trial-and-error method. The hyperbolic tangent was selected as the activation function. The predicted temperature, relative humidity, and surface temperature satisfied the ASHRAE guidelines (Table 2 and Figure 7).

**Table 2.** Test of model performance by ASHRAE guideline

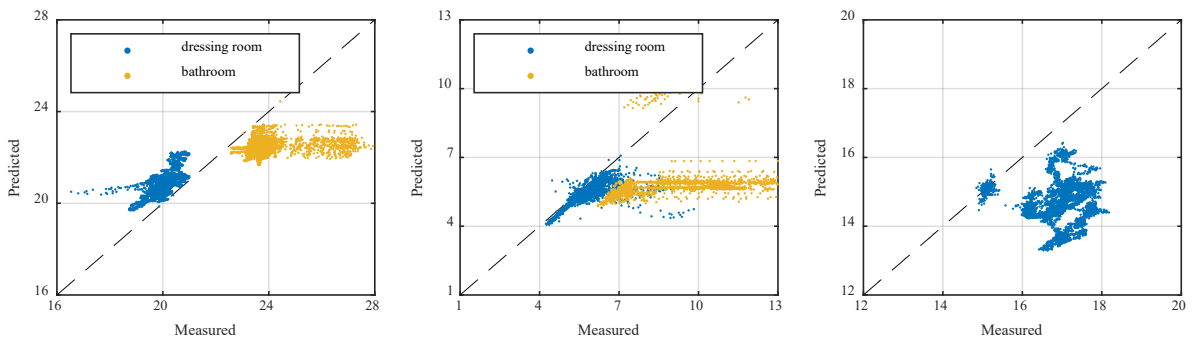
		Temperature	Humidity	Surface temperature
Physics-based	CVRMSE(%)	6.13	69.63	16.53
	MBE(%)	-0.02	-0.44	-0.15
Grey-box	CMRMSE(%)	3.45	52.83	14.29
	MBE(%)	-0.01	-0.33	-0.13
Data-driven	CVRMSE(%)	1.98	7.12	0.98
	MBE(%)	0.02	0.01	-0.01

## 4. Conclusion

In this paper, three models were developed: (1) physics-based model, (2) grey-box model, and (3) data-driven model. Despite the generalization performance and analytical explanation of the physics-based model based on the first principles, it exhibits the lowest accuracy. This is because the field measurement cannot be designed for an ideal situation wherein everything is under control, as in an experimental chamber. To determine the unknown parameters in the physics-based model, the authors used the

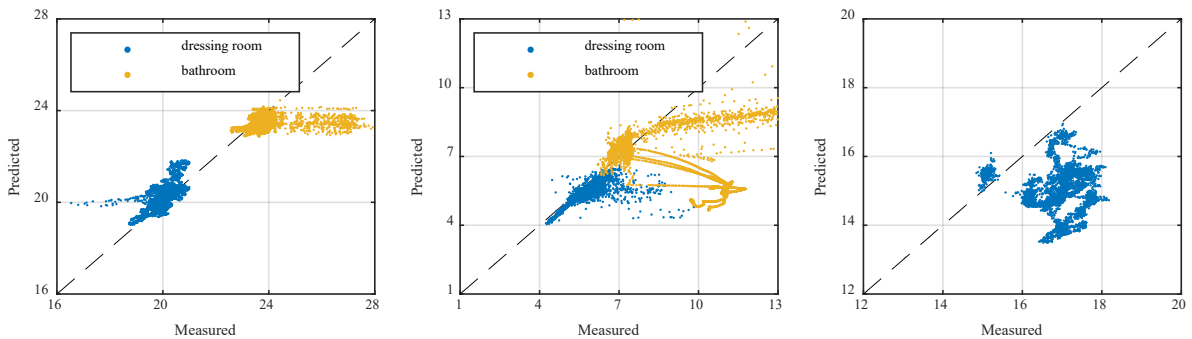


**Figure 4.** Structure of ANN model prediction



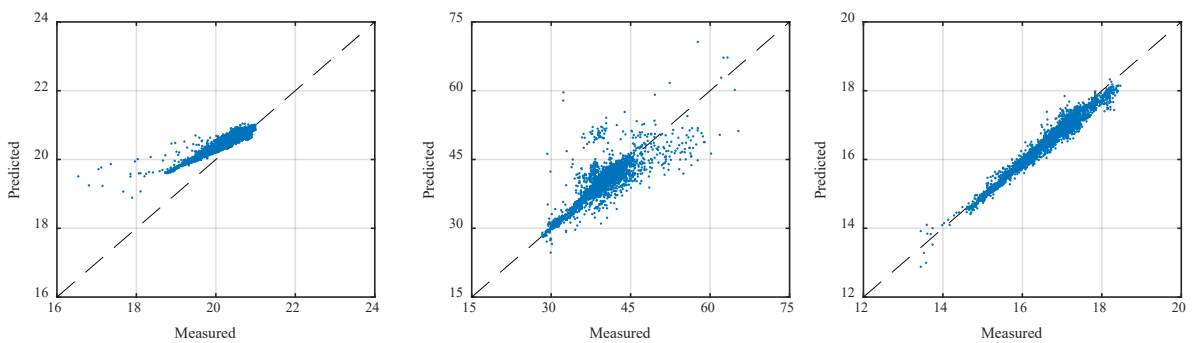
(a) Temperature (°C) (b) Absolute humidity (g/kg) (c) Surface temperature (°C)

**Figure 5.** Physics-based model predicted values vs. measured data



(a) Temperature (°C) (b) Absolute humidity (g/kg) (c) Surface temperature (°C)

**Figure 6.** Grey-box model predicted values vs. measured data



(a) Temperature (°C) (b) Relative humidity (%) (c) Surface temperature (°C)

**Figure 7.** ANN model predicted values vs. measured data

parameter estimation technique and used it as a grey-box model. Although the average prediction performance has been improved (CVRMSE: 30.76% → 23.52%; MBE: -0.20% → -0.16%), the absolute humidity still did not meet the criteria owing to uncertain factors. Using measured data that was once used for parameter estimation, the ANN model was then developed and was found to exhibit the highest accuracy in all respects. Although the largest number of input variables was used for training, it does not require data that is difficult to measure, such as the convective heat transfer coefficient or barometric pressure. Furthermore, there is no requirement for “calculating” the absolute humidity at each node based on the temperature and relative humidity using an additional step. Only the original data is used to define the relation between the input and output variables in the data-driven approach. Although the application of this simulation model is limited in other buildings, with respect to MPC, the data-driven model can be the most economic option with lighter work required for model development and a higher prediction performance in condensation prediction for the specific area under consideration. Based on this prediction model, the optimal control strategies for minimizing the surface condensation will be tested in a future study.

### Acknowledgement

This research was supported by a grant(19RERP-B082204-06) from Residential Environment Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

### Reference

- [1] TenWolde, A 1994 Ventilation, humidity, and condensation in manufactured houses during winter *ASHRAE Tran* **100** 103-115
- [2] R. Becker 1984 Condensation and mould growth in dwellings—parametric and field study *Build Environ* **19(4)** 243-250
- [3] Huibo Zhang and Hiroshi Yoshino 2010 Analysis of indoor humidity environment in Chinese residential buildings *Build Environ* **45(10)** 2132-2140
- [4] Shahryar Motakef and Mäher A. El-Masri 1986 Simultaneous heat and mass transfer with phase change in a porous slab *Int J Heat Mass Tran* **29(10)** 1503-1512
- [5] Tao Lu and Martti Viljanen 2009 Prediction of indoor temperature and relative humidity using neural network models: Model comparison *Neural Comput Appl* **18(4)** 345- 357
- [6] Leopold Mba, Pierre Meukam and Alexis Kemajou 2016 Application of artificial neural network for predicting hourly indoor air temperature and relative humidity in modern building in humid region *Energ Buildings* **121** 32-42
- [7] Abdul Afram and Farokh Janabi-Sharifi 2014 Theory and applications of HVAC control systems – A review of model predictive control (MPC) *Build Environ* **72** 343-355
- [8] Jing Liu, Hiroyoshi Aizawa and Hiroshi Yoshino 2004 CFD prediction of surface condensation on walls and its experimental validation *Build Environ* **39(8)** 905-911
- [9] Sten de Wit and Godfried Augenbroe 2002 Analysis of uncertainty in building design evaluations and its implications *Energ Buildings* **34(9)** 951-958
- [10] Jerzy Wyrwal and Andrzej Marynowicz 2002 Vapour condensation and moisture accumulation in porous building wall *Build Environ* **37(3)** 313-318
- [11] Qiang Zhou, Shengwei Wang, Xinhua Xu and Fu Xiao 2008 A grey-box model of next-day building thermal load prediction for energy-efficient control *Int J Energ Res* **32(15)** 1418-1431
- [12] Cheol-Soo Park, Godfried Augenbroe, Tahar Messadi, Mate Thitisawat and Nader Sadegh 2004 Calibration of a lumped simulation model for double-skin façade systems *Energ Buildings* **36(11)** 1117-1130
- [13] M. De Moor and D. Berckmans 1996 Building a grey box model to model the energy and mass transfer in an imperfectly mixed fluid by using experimental data *Math Comput simulat* **42(2-3)** 233-244
- [14] ASHRAE 2014 ASHRAE Guideline 14-2014 for Measurement of Energy and Demand Saving Atlanta: American Society of Heating, Refrigerating and Air-conditioning Engineers, Inc.