A long short-term memory-based prediction model for gas concentrations and thermal environment conditions in a closed pig building

Q. Xie[1], J.-Q. Ni[2], J. Bao[3,4,5*] and P. Zheng[1]
[1]*College of Electrical and Information, Northeast Agricultural University, Harbin 150030, China*
[2]*Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907, USA*
[3]*College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China*
[4]*The Key Laboratory of Swine Facilities Engineering, Ministry of Agriculture and Rural Affairs, Harbin 150030, China*
[5]*Engineering Research Center of Pig Intelligent Breeding and Farming in Northern Cold Region, Ministry of Education, Harbin 150030, China*
*Correspondence author: Jun Bao, jbao@neau.edu.cn

Abstract

To improve the indoor environment conditions in pig buildings, an innovative prediction model with three layers using long short-term memory (LSTM) was developed to forecast indoor environment conditions including five factors of the gas concentrations of ammonia (NH_3), carbon dioxygen (CO_2) and hydrogen sulfide (H_2S), and thermal environment conditions of air temperature (PLS T) and relative humidity (PLS RH). In the LSTM model, there are 5 inputs and outputs representing for the five environmental factors in history and the next time step, respectively. The auxiliary influenced factors with high correlations with the outputs were selected to attend the calculation in fully connected layer. 53 days collected data in wintertime and summertime were used for the model training and testing. It was shown that the model with hyper-parameters of 400 neurons, 1 time lag, and 0.008 learning rate has optimal performances both in summertime and wintertime. Compared with other two different types of models of ANFIS and MLR, the LSTM model has the best performances of RMSE, MAPE and R^2 for the concentrations of NH_3 in wintertime, H_2S in summertime, PLS T and PLS RH both in summertime and wintertime.

Keywords: thermal environment conditions, gas concentrations, environment control, pig building, long short-term memory.

Introduction

In intensive livestock productions, the indoor environment conditions are key factors for livestock (Hay, Morton et al., 2016) because of their great influence on the animal's growth, welfare, health, feed-to-meat conversion, and reproduction. The major environmental factors in pig living space include air temperature and humidity, gases of ammonia (NH_3), carbon dioxide (CO_2) and hydrogen sulfide (H_2S).

To improve the environment conditions and meet the specific environmental needs of pigs in different stages and areas, in the past decades, some field experimentation and simulation-or-modelling-based research on indoor environment control in animal buildings were conducted. The field experiments were carried out on ventilations (Kim, Ko et al., 2007, Chen, Lim et al., 2014, Ni, Kaelin et al., 2016), temperature and humidity (Samer, Loebsin et al., 2011, Hempel, König et al., 2018), gases concentration and emission (Saha, Zhang et al., 2010, Zong, Zhang et al., 2014, Zong, Li et al., 2015), etc. However, in practice, due to the drawback caused by the limitations of practical measurement sensors' position and number, time delay of the control devices, etc., it is not easy to achieve the desired environment conditions in time.

As recent evolutions in sensor technology and big data have provided continuously expanding resources for data collection. Many data-driven prediction methods, like linear or nonlinear method, numerical method like Computational fluid dynamics (CFD), especially neuron networks such as Artificial Neural Network (ANN), Elman, Backpropagation (BP) and Adaptive-Network-based Fuzzy Inference Systems (ANFIS) have been used by scholars to forecast the environment conditions in animal buildings. However, as large volume of input data increased, most of the neural networks in these models were rarely based on multilayer neural networks which were not specifically designed for time series model.

Long and short-Term memory (LSTM) is a deep learning method that performs very well in dealing with the problems of the time series data (Gers, Schmidhuber et al., 2000). LSTM provides new ideas for solving the problems of predictions in air pollution, building energy, traffic flow and travelling time, recognition and classifications. Good performances of these predictions for large data volume using LSTM have exhibited the full potential of artificial neural networks. However, to the best of the authors' knowledge, attempts have not yet been found to use LSTM algorithm considering other influenced factors to predict multiple time series of indoor environment factors, such as air temperature, humidity, and harmful gas concentrations of NH_3, CO_2 and H_2S, for livestock building in research papers.

The object of this study is to address the aforementioned limitations and propose a novel model using the deep learning algorithm of LSTM to improve the prediction of multiple indoor environment factors and provide optimal control strategies for the indoor environment management in pig buildings.

Pig building and measurement data

Building description

Measurement data in this study were obtained in a 12-room swine environmental research building at Purdue University, USA. Each room is consisted of a 11.0 m × 6.1 m × 2.7 m (L × W × H) pig living space (PLS) and two 1.8-m deep manure pits (Liu, Ni et al., 2017). The PLS and the manure pits are separated with a slatted concrete floor (Figure 1). The PLS has a capacity of housing 60 finishing pigs in two rows of 3–6 pens.

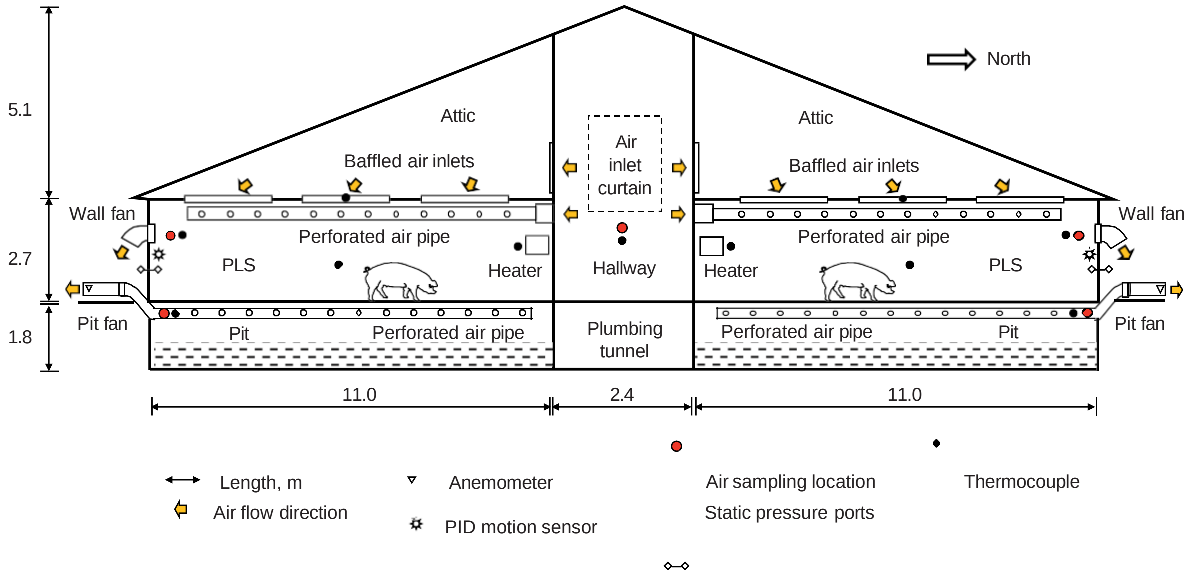

Figure 1: Cross-sectional east-side view of the swine building.

A propane heater (Model Guardian 60, L.B. White Co. Onalaska, WI, USA) was installed in each room for heating when the room temperature was too low for piglets in winter. Four ventilation fans: two pit fans of

250-mm diameter and two wall fans, one of 356-mm and another of 508-mm diameters were equipped for each room. Heating and ventilation were independently controlled with a Fancom controller (Model FCTC, Vancom BV, Panningen, The Netherlands) in each room. Fresh air entered into the building from two air inlets on top of the east and west doors, supplied to each room from ceiling and hallway inlets.

The building was equipped with a continuous environmental monitoring system to measure temperature, relative humidity, ventilation, and concentrations of different gases.

Data selection

Several research projects on animal environment have been conducted in the past years in the swine building. Data from Room 2 in summertime from August 28 to September 30, 2015 and from Room 4 in wintertime from December 24, 2015 to January 13, 2016 were selected as the input data and test data of the prediction model in this study. During these two periods, the pig numbers in Room 2 and Room 4 were 64–146 and 60, respectively; the total pig weights were 1942 – 4157 kg in Room 2 and 6212 – 7719 kg in Room 4.

Measurement data, including temperature, relative humidity, ventilation, gas concentrations, and pig activities, at 1-min resolution were averaged to hourly data and were used for model training, and model test and validation.

LSTM Model

Networks of LSTM have a memory block connected through layers that consists of a set of memory blocks, and at the same time forgetting irrelevant information, which mimics the way in which human selectively remember details from the past.

Parameter selection: With the consideration of the outputs have different correlations with the factors, two prediction categories are established: (1) predictions for gas concentrations of NH3, CO2 and H2S; (2) predictions for indoor thermal environment of PLS T and PLS RH. The factors of Pit NH3, Pit CO2, Pit H2S, PLS T, PLS RH, Out RH, Out T, Act, Wt and Vent are selected as the auxiliary input variables for category (1) to attend the forecasting of the room gas concentrations of NH3, CO2 and H2S; the factors of Out RH, Out T, Act, Wt and Vent were selected as the auxiliary input variables for category (2) to attend the prediction of the indoor thermal environment of PLS T and PLS RH.

Model structure: The model consists of three layers: sequence input layer, hidden layer and output layer. The hidden layer includes LSTM layer, fully connected layer, and regression layer (Figure 2).

The history sequence data of concentrations of NH_3, CO_2 and H_2S, PLS T and PLS RH were transformed into the input times series named x_1, x_2, ……, x_5, the normalized auxiliary data as the influencing factors on each input variables of x_1, x_2, ……, x_5 attending the calculation of the fully connected layer in the hidden layer. The dimensions of the input time series are 5×t, t is the length of the time series.

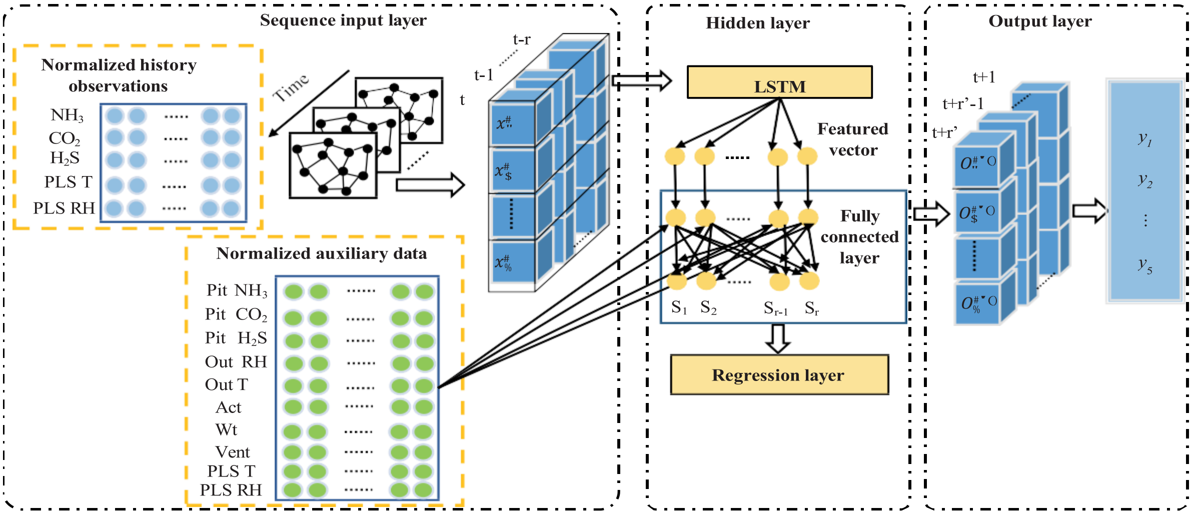

Figure 2: Model structure.

Three layers of LSTM, fully connected and regression are constructed in the hidden layer to achieve a better learning effect. The LSTM layer learns long-term dependencies between time steps in time series and sequence data to performs some interactions that help to improve gradient flow over long sequences.

The output layer is responsible for the demoralization of the output values O_1, O_2, ……, O_5 getting from the hidden layer into predictions y_1, y_2, ……, y_5 of the same format as the observations representing for predictions of the concentrations of NH_3, CO_2 and H_2S, PLS T and PLS RH.

The same model structure is used for the two prediction categories of (1) and (2), the difference lies in the auxiliary factors according to different correlations with the outputs.

Model development: The time sequence data are divided into two sets: training data set and testing data set. The training data are used for LSTM model training, and the testing data are used for testing the performance of the model. In this study, the total 1320 observations sampled from Aug. 28 to Sept. 30, 2015 (representing for summertime) and Dec. 24 to 31, 2015 and Jan. 1 to 13, 2016 (representing for wintertime), were divided into four data sets according to different seasons. The sampled data in the last two days were used as the testing data, and others were used for training data.

The computation of data training and testing for the LSTM model was conducted using the Matlab.In the LSTM layer, the initial network parameters are set: the number of hidden neurons is 200, the state activation function is hyperbolic tangent (tanh), the gate activation function is sigmod, the training method is Adam (adaptive moment estimation) optimizer, the maximum epochs are 250, the initial learning rate is 0.005. After training, the LSTM prediction model for multiple environment factors will be obtained.

Model evaluation: The performances of the LSTM model were evaluated with statistical analysis, such as root mean squared error (RMSE), mean absolute percentage error (MAPE) and coefficient of determination (R2).

Results and discussion

Prediction result of the LSTM model

The prediction model with learning rate of 0.008, neurons of 400, and time lag of 1 (for PLS T, PLS RH and concentration of NH_3 and CO_2) and 2 (for H_2S) was built, in this paper, to increase the model's accuracy and useful features learning.

Comparison with different models

In order to evaluate the availability of the proposed deep learning algorithm of LSTM model, other two methods of adaptive neural fuzzy inference system (ANFIS) (Takagi and Sugeno 1985) and multiple linear regression (MLR) are also employed to forecast PLS temperature, PLS relative humidity, concentrations of NH_3, CO_2 and H_2S using the same dataset.

The overall variation trends of the predictions with the three methods both in winter and summer are consistent with the observations (Figure 3). Also, they are coincided with the measured data that the concentrations of NH_3, CO_2 and H_2S in summer are lower than those in winter (Figure 3 A, B and C), and the PLS T and PLS RH in summer are higher than those in winter (Figure 3 D and E).

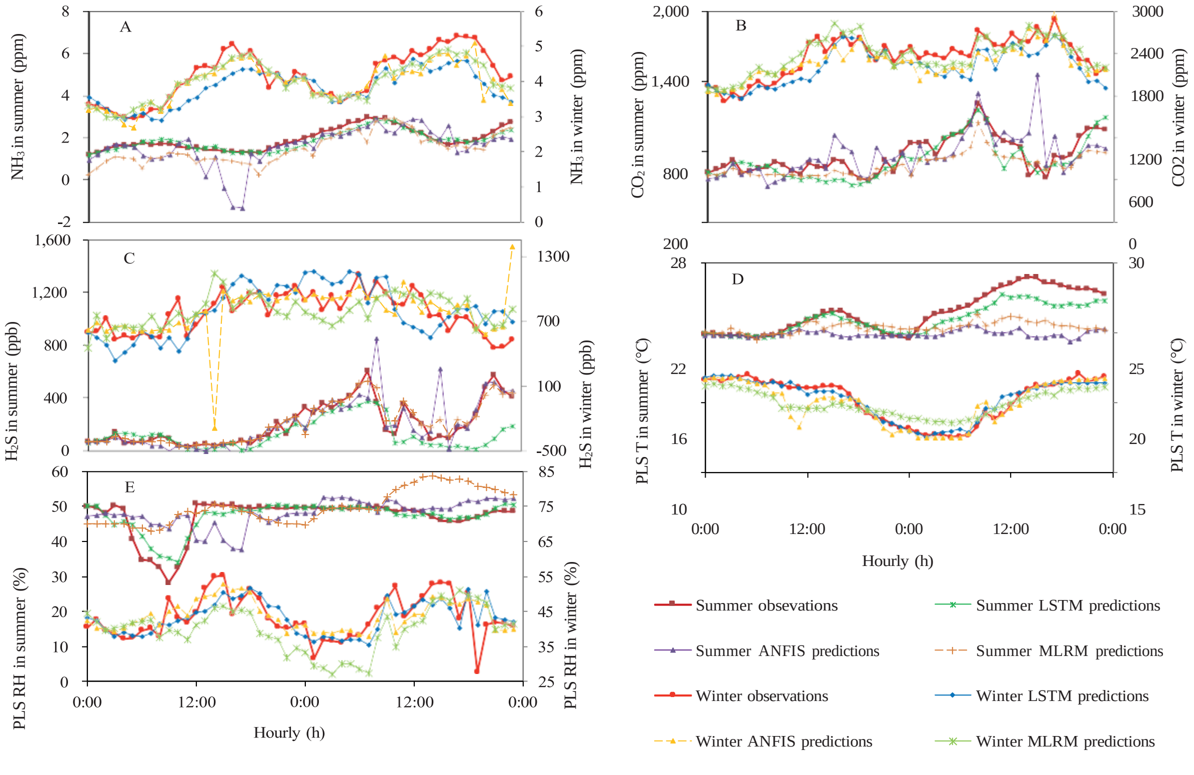

Figure 3: Comparisons between the observations and predictions with method of LSTM, ANFIS and MLR in summertime and wintertime. A, B and C: Concentrations of NH_3, CO_2 and H_2S, respectively; D: PLS temperature; E: PLS relative humidity.

The performances of RMSE, MAPE and R^2 using the LSTM methods for the predictions of NH_3 in summertime are 0.1865, 7.6565 and 0.8587, respectively, as well as those of H_2S in wintertime of 57.3827, 16.9303 and

0.6126, PLS T of 0.8473, 2.5040 and 0.7034 (in summertime), and 0.3621, 1.3917 and 0.9377 (in wintertime), and PLS RH of 2.3460, 3.7687 and 0.7467 (in summertime), and 4.2471, 7.2806 and 0.5419 (in wintertime), which are much better than those using the methods of ANFIS and MLR.

Therefore, from the results of prediction curve and the performances that compared above, the predictions for gas concentrations of NH_3, CO_2, and H_2S, PLS T and PLS RH using the LSTM method are more optimal than those using other two methods of the ANFIS and MLR.

Conclusions

The conclusions were drawn in this study:

(1) It was found that the LSTM model with the number of neurons of 400, time lag of 1, and learning rate of 0.008, could achieve better performances of RMSE, MAPE and R^2 both in summertime and in wintertime. The LSTM model for the concentrations of NH_3 in wintertime, H_2S in summertime, PLS T and PLS RH both in summertime and wintertime had better output performances of RMSE,

(2) MAPE and R^2 compared with those using the methods of ANFIS and MLR, at the same time, it also demonstrated that the deep learning method of LSTM is more suitable for the predictions of time series.

(3) The average errors between predictions and observations during the two periods of summertime and wintertime for concentrations of NH_3, CO_2 and H_2S, PLS T and PLS RH were 0.197 ppm, 92.7855 ppm, 27.6414 ppm, 0.2579 °C, and 0.0361%, respectively.

Acknowledgments

This work was by the project of the National Natural Science Foundation of China (NSFC) (32072787); the project of the Postdoctoral Science Foundation of Heilongjiang Province (LBH-Q21070), China; the Key Laboratory of Swine Facilities Engineering, Ministry of Agriculture, China. This work was also supported by the USDA National Institute of Food and Agriculture Hatch project, USA (1011562). The authors acknowledge the support by the researchers and managers at the Animal Science Department, Purdue University, USA during the building environment measurement, and the help by all others for this study.

References

Chen, L., Lim, T.-T., Jin, Y., Heber, A.J., Ni, J.-Q., Cortus, E.L., and Kilic, I. (2014) Ventilation rate measurements at a mechanically-ventilated pig finishing quad barn. *Biosystems Engineering* 121, 96-104.

Gers, F.A., Schmidhuber, J., and Cummins, F. (2000) Learning to forget: Continual prediction with LSTM. *Neural Computation* 12(10), 2451-2471.

Han, J. (2018) *Dehumidification technology evaluation and moisture balance modelling for greenhouse humidity control* University of Saskatchewan].

Hay, K., Morton, J., Mahony, T., Clements, A., and Barnes, T. (2016) Associations between animal characteristic and environmental risk factors and bovine respiratory disease in australian feedlot cattle. *Preventive Veterinary Medicine* 125, 66-74.

Hempel, S., König, M., Menz, C., Janke, D., Amon, B., Banhazi, T.M., Estellés, F., and Amon, T. (2018) Uncertainty in the measurement of indoor temperature and humidity in naturally ventilated dairy buildings as influenced by measurement technique and data variability. *Biosystems Engineering* 166, 58-75.

Kim, K.Y., Ko, H.J., Kim, H.T., Kim, Y.S., Roh, Y.M., and Kim, C.N. (2007) Effect of ventilation rate on gradient of aerial contaminants in the confinement pig building. *Environmental Research* 103(3), 352-357.

Liu, S., Ni, J.-Q., Radcliffe, J.S., and Vonderohe, C.E. (2017) Mitigation of ammonia emissions from pig production using reduced dietary crude protein with amino acid supplementation. *Bioresource Technology* 233, 200- 208.

Ni, J.-Q., Kaelin, D., Lopes, I.M., Liu, S., Diehl, C.A., and Zong, C. (2016) Design and performance of a direct and continuous ventilation measurement system for variable-speed pit fans in a pig building. *Biosystems Engineering* 147, 151-161.

Saha, C.K., Zhang, G., Kai, P., and Bjerg, B. (2010) Effects of a partial pit ventilation system on indoor air quality and ammonia emission from a fattening pig room. *Biosystems Engineering* 105(3), 279-287.

Samer, M., Loebsin, C., Fiedler, M., Ammon, C., Berg, W., Sanftleben, P., and Brunsch, R. (2011) Heat balance and tracer gas technique for airflow rates measurement and gaseous emissions quantification in naturally ventilated livestock buildings. *Energy and Buildings* 43(12), 3718-3728.

Takagi, T., and Sugeno, M. (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*(1), 116-132.

Zong, C., Li, H., and Zhang, G. (2015) Ammonia and greenhouse gas emissions from fattening pig house with two types of partial pit ventilation systems. *Agriculture, Ecosystems and Environment* 208, 94-105.

Zong, C., Zhang, G., Feng, Y., and Ni, J.-Q. (2014) Carbon dioxide production from a fattening pig building with partial pit ventilation system. *Biosystems Engineering* 126, 56-68.