



HAL
open science

Development of deep learning artificial neural networks models to predict temperature and power demand variation for demand response application in cold storage

Hong-Minh Hoang, Mahdjouba Akerma, Nedra Mellouli, Alan Le Montagner, Denis Leducq, Anthony Delahaye

► To cite this version:

Hong-Minh Hoang, Mahdjouba Akerma, Nedra Mellouli, Alan Le Montagner, Denis Leducq, et al.. Development of deep learning artificial neural networks models to predict temperature and power demand variation for demand response application in cold storage. *International Journal of Refrigeration*, 2021, 131, pp.857-873. 10.1016/j.ijrefrig.2021.07.029 . hal-03321382

HAL Id: hal-03321382

<https://hal.inrae.fr/hal-03321382>

Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Development of deep learning artificial neural networks models to predict temperature and power demand variation for demand response application in cold storage

H.M. Hoang^a✉, M. Akerma^a, N.Mellouli^b, A. Le Montagner^a, D. Leducq^a, A. Delahaye^a

^a Université Paris-Saclay, INRAE, FRISE, 92761, Antony, France

^b LIASD EA4383, IUT de Montreuil, Université Paris 8, Vincennes Saint-Denis,
France

ABSTRACT

Food warehouses and cold rooms have a significant potential for Demand Response (DR) application (stopping or reducing the power of fans and compressors of the refrigeration system) due to thermal inertia of food products. However, as air and food temperature might increase beyond acceptable limits during DR periods, DR needs to be carefully applied in order to respect the food temperature regulation and to maintain quality and safety of the products. It is thus important to predict the system behaviour in case of DR application in order to evaluate its potential impacts and to decide if DR can be performed or not. Four deep learning artificial neural networks (ANN) models, traditional Long Short-Term Memory LSTM, stacked LSTM, bidirectional LSTM and convolutional LSTM, were developed to predict future temperature and power demand perturbations due to the application of DR in cold storage. The aims of this work are: first, to assess the performance of those models in predicting the system behaviours, in particular the sudden variations during and after DR applications, and second, to identify the impact of data availability (number of sensors, their positions) and data characteristic (quality, quantity and DR patterns) on the prediction performance. The results have shown the high potential of deep learning ANN models in supporting DR application in cold storage.

KEYWORDS

Demand response, cold storage, electric energy consumption, Deep Learning, long short-term memory, neural network

✉Corresponding author: Tel: 33 1 40 96 65 02 Fax: 33 1 40 96 60 75

e-mail: hong-minh.hoang@inrae.fr

ABBREVIATION

ANN	Artificial neural networks
ARIMA	Autoregressive integrated moving average
ARMA	Autoregressive moving average
BPNN	Back propagation neural networks
CNN	Convolutional neural networks
DNN	Deep neural networks
DR	Demand response
DT	Decision tree
GHG	Greenhouse Gases Emissions
GRNN	General regression neural networks
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
RBNN	Radial basis neural networks
RES	Renewable energy sources
RNN	Recurrent neural networks
SVM	Support vector machine
VARMA	Vector autoregressive moving-average

NOMENCLATURE

b	Bias vector
c_t	Cell state vector
\tilde{c}_t	Cell input activation vector
Comp	Binary input: compressor ON (Comp =1); OFF (Comp =0)
Def	Binary input: defrost ON (Def = 1); OFF (Def = 0)
DR	Binary input: in DR period (DR = 1); not in DR period (DR = 0)

$\delta_{i,def}$	Time elapsed since the last defrost period, s
$\delta_{i,DR}$	Time elapsed since the last DR period, s
f_t	Cell forget gate's activation vector
h_t	Cell hidden state (output) vector
H	Prediction horizon
i_t	Cell input gate's activation vector
m	Number of observations
N_e	Number of external air temperature sensors
N_r	Number of return air temperature sensors
o_t	Cell output gate's activation vector
P	Compressor demand power, W
σ_A	Activation function
T_{ext}	Air temperature in the external cell, °C
T_p	Product temperature, °C
T_r	Return air temperature, °C
T_s	Supply air temperature, °C
U	Weight matrices
x_t	Model inputs
y_t	Model outputs
\hat{y}_t	Estimation of y
W	Weight matrices

1. Introduction

In the current context of energy transition, there is a growing use of renewable energy sources (RES) in power systems: 27 % share of RES is expected in gross final energy demand in 2030 in Europe [1]. Because of the intermittent nature of solar and wind energy, two main sources of renewable energy, one of the main concerns of energy management nowadays is how to balance energy demand and supply. Demand response (DR), one of the energy management strategies, is considered as an elegant solution to increase the flexibility of end-user demands by changing their electricity profiles from their

usual consumption patterns [2]. For example, when there is a diminution of electricity generation by RES due to environmental conditions, end-users such as industrial sites might reduce their power demand to respond to this change of electricity supply. Thus, DR is often viewed as a way to change the end-users' role from a passive to an active way. Other benefits of DR application in terms of financial incentives and consumption reductions can also be expected [3, 4].

With about 600 million cubic meters of storage volume worldwide [5], food warehouses and cold rooms have an important potential for DR application due to the thermal inertia of food products. Different DR strategies can be applied for refrigeration applications: increase the setpoint temperature [6, 7]; stopping or reducing the power of evaporator fans; stopping or reducing the power of compressors of the refrigeration system [8], etc. Cold storages, by their thermal inertia due to stored products, could be considered as "thermal batteries". In a certain manner, they replace the refrigeration system and limit the temperature rise during DR application. The more important the volume is, the more thermal energy is stored and thus the more important potential for DR application. DR application in frozen food storage presents a stronger potential than in chilled food storage, the chilled products being usually more impacted by temperature variations than frozen ones. However, even with frozen products, DR periods need to be carefully applied to respect the food temperature regulation and maintain their quality and safety. Indeed, air and product temperatures increase during each DR application and then decrease when the refrigeration system returns to its normal functioning. Many studies [9-11] have shown that frozen food quality is impacted by both temperature storage level and temperature variation amplitude. Industrials will refuse DR applications if the risks of product quality and safety degradation are too strong. Besides the risk related to product, DR might generate negative impacts on power demand: with an increase of power demand when restarting the compressor (postponing power effect) or over-consumption of energy for the recovery of the temperature level before DR (postponing consumption effect) [8].

Actually, in the near future the cold chain could be forced (or it could be an opportunity for it) to be involved with Demand Side Management and/or Demand Response (DR). It is therefore of paramount importance to understand the possible behaviour of refrigerated spaces and refrigerated items when refrigerated systems undergo DR protocols. Indeed, for the development of DR in cold storage, it is primordial to predict the system behaviour in case of DR application in order to evaluate its potential impacts and to decide whether DR could be performed or not. However, the numerical prediction of the change due to DR application in terms of temperature and energy consumption might face following challenges:

- a) Multi-factor dependency: many parameters such as load percentage, setpoint temperature, DR duration, outside temperature, etc. might exert influences on system behaviour [8];
- b) Abrupt variation: the turning off of the compressor might generate a sudden air temperature rise which is difficult to be predicted by numerical models: 10°C of air temperature rise was observed in an empty cold room after 1 h of compressor off [8].
- c) Data gathering: few studies were done in this subject so that data for model development and validation are lacking.

It is emphasized that the difficulty in cold storage system modelling is also related to the dynamic behaviours of refrigeration systems (compressor on-off cycles, defrost, Fig. 1). The air flow pattern inside the cold room is complex and depends on the product layout [12]. Moreover, DR implementation adds a completely new behaviour and new DR related data are needed. The behaviours of product and air are very different during and after DR application when the compressor and fan power might be reduced or turned off. As the air was warmed up because of thermal loss

through the wall, it is then the product that cools down the air and slows down the rate of the temperature increase. While a conventional lumped thermal capacity model can be applied for a small storage room at almost constant conditions, it is not easy to develop this kind of model for DR application. Indeed, thermal models need deep information regarding physical parameters: wall geometry and properties, insulation thickness, refrigeration machine parameters (compressor and fan powers, set-point temperature, refrigerant, defrost setting...), product mass and initial temperature...The development of dynamic thermal models for warehouses is even more challenging with regard to the wide variety of warehouses (sizes, fan placement...); the frequently changing load configuration (mass and disposal) that modifies the air flow pattern; the complexity of staff interactions;...Thus, there is growing scientific interest in developing alternative modelling approaches other than thermal modelling.

In recent years, the rapidly increased availability of measurement data of physical systems, the reduction of cost of sensors and the possibility of storage and process important data quantity have led to the development of many data driven methods for modelling and predicting dynamic behaviours [13]. In particular, deep neural networks (DNNs which are artificial neural networks - ANN with multiple layers between the input and output layers) are at the cutting edge of data-driven methods. DNNs reach not only superior performance for tasks such as image classification, but they have also proven to be effective in predicting the future state of dynamical systems [14]. In this work, from a review of data driven approaches, recurrent deep learning models were selected for DR forecasting using data collected in a laboratory cold room. These predictive models could be easily extended for application in cold warehouses using transfer learning [15], a technique developed for DNNs for the transfer of knowledge learnt from one condition (cold room) to similar conditions (targeted cold warehouses).

The paper is structured as follows. The related work is presented in section 2 followed by the contribution of this work (section 3). Then in section 4 the experimental setup, data acquisition and collection are described while the section 5 is focused on the prediction problem formulation and model development. The results of the developed models are discussed in section 6.

2. Related works

Three types of modelling approaches have been developed for predicting energy system behaviour: data driven, physics-based and "Grey Box" [16]. Using equations describing physical phenomena that occur in the system, physics-based (or White Box) approach requires detailed knowledge on system characteristics. For example, if this approach is applied to a warehouse, detail information for the calculation of heat exchanges inside the warehouse (thermophysical properties of the products, warehouse geometry, wall and insulation composition, refrigeration machine regulation ...) are needed [17-19]. Due to this high complexity and when the main objective is the final result obtained at the output of the system independently of internal operations, it might be interesting to examine the use of data driven (or Black Box) models. More details on this approach will be discussed further in this section. The third type, Grey Box, is developed from the coupling of White Box and Black Box approaches. It uses a hybrid model based on a simplified physical representation of the system structure, the parameters of this model being identified using data driven approach [20-22].

The modelling of the dynamic behaviour of cold storage system using data driven models can be approached as the study of time series, defined as a sequence of discrete time data. Time series consists of indexed data points, measured typically at successive times, spaced at often uniform time

intervals. A time series forecasting model predicts future values based on known past events (recent observations). Conventional time series prediction methods commonly use a moving average model in which the output variable depends linearly on the current and various past values. These models can be autoregressive (ARMA) [23], integrated autoregressive (ARIMA) [24, 25] or vector autoregressive (VARMA) [26] in order to reduce data. They have been used successfully to predict electrical load and energy consumption in building and industries. However such methods lack flexibility and generalizability as they must process all available data in order to extract the model parameters that best fit the new data. These methods are not suitable face to massive data, data with abrupt variations and real-time series forecasting for decision support as in DR application.

In this context, non-linear models which are more flexible and more generalizable have been developed. Recently, artificial neural networks (ANN) have shown great potential for application in energy system [27, 28]. ANNs have developed into a large family of techniques from simple networks with 3 layers (input, hidden and output) to complex multiple layers. Li *et al.* [29] used 3 simple ANNs (back propagation BPNN, radial basis function RBNN and general regression GRNN) and support vector machine (SVM) for the prediction of hourly cooling load in the building; better accuracy and generalization were obtained with SVM and GRNN. When facing a complex problem, an intuitive solution is to add more layers to form “deep learning” network (DNN) and increase its learning capacity and to capture detailed information. In particular, Long Short-Term Memory (LSTM) [30] is a class of recurrent neural networks (RNNs) [31] capable of preserving both short-term (recent variations, actual tendencies of data) and long-term information (periodicity, recurrent events or not). LSTM which is designed for sequential data has particularly shown promising results for time series prediction: LSTM deep learning models were used successfully in forecasting indoor temperature with better performance than SVM, BPNN and decision tree (DT) models [32]. It is worth to be noted that LSTM corresponds to a family of ANN models. For example, LSTM can be combined with convolutional (CNN) network to form CNN-LSTM model. Kim and Cho [33] have shown the potential of CNN-LSTM deep learning model in extracting complex features (CNN attribute) and temporal information (LSTM attribute) of energy consumption in residential building. Another variation of LSTM, bidirectional LSTM (Bi-LSTM) consists of connecting LSTM layer with both past and future data. Higher performance of Bi-LSTM model than traditional LSTM model was observed by Das *et al.* [34], in particular with longer prediction horizon.

Few works have applied deep learning ANN models in energy systems for DR application: predict prices and energy demand evolution in smart grid [35], forecast load demand in residential building [36]. Recently, deep learning models have been applied to refrigeration domain. Onoufriou *et al.* [37] used RNN models to predict air temperature evolution during defrost. This study highlighted that slight variations in data availability and consistency can have a significant effect on the learning performance.

While classical physical models might encounter difficulties in modelling refrigeration systems, characterized by the non-linearity and the coupling of several temporal parameters, in particular during disturbances such as DR periods, DNN methods, due to their ability to adjust and self-learning, can therefore be very promising in responding to this type of issues offering a new way to solve complex problems. LSTM ability in mapping complex non-linear relationships, has succeeded in several problems such as planning, control, analysis and design. The literature has demonstrated their superior capability over conventional methods, their main advantage being the high potential to model non-linear processes [38]. In the rest of this paper, the study was focused on cold room (source) and it's assumed that various physical behaviours, in particular variations in air and product temperatures due

to DR applications, defrost and on/off compressor working cycles, are close to those of a cold warehouse (target). In future works, the LSTM models pre-trained on the source dataset of a cold room will be fine-tuned on the target dataset of a cold warehouse.

3. Contribution of this work

In view of the potential of LSTM deep learning models, 4 LSTM models (traditional LSTM, Convolutional LSTM, Stacked LSTM and Bidirectional LSTM) have been developed by our team for DR application in cold storage.

The two main contributions of this work are:

a) Prediction performance assessment of 4 LSTM models for DR application in cold storage

Particular attention is given to LSTM deep learning models to determine the extent to which they are capable of predicting DR behaviour in terms of temperature and power demand variations. The performance of 4 LSTM models (traditional LSTM, Convolutional LSTM, Stacked LSTM and Bidirectional LSTM) was evaluated over different prediction horizons to identify:

- which models are the most suitable to represent specific data related to DR application
- which models are the most sensitive to the variations of temperature and power demand
- which models are the most robust to signal noises?

b) Evaluation of the influence of data availability and characteristics on model prediction performance

As previously stated, data gathering in industrial cold storage is a challenging task, in particular data related to DR applications. This is why experimental data collected in a laboratory cold room [8] will be used in this work. Many questions have been raised about the needed data quality (how temperature and power demand variations during DR can be well depicted? presence of noise?) and quantity (number of points?). It is often difficult to have as many sensors installed in an industrial warehouse as wished. If the number of sensors is limited, which are their best positions inside a cold room/warehouse? In order to answer those questions:

- The impact of data availability on the prediction will be analysed: by using 3 configurations of inputs and outputs from different combinations of sensors (Cf. 5.1.2. Configurations of inputs and outputs).
- The impact of data characteristics (quality, quantity and DR patterns) on the prediction will be evaluated by using 2 datasets (Cf. Table 1).

4. Experimental setup - Data acquisition and collection

4.1. Cold room, sensors and measurement protocol

Experimental data of DR applications were obtained in a frozen cold room (dimensions L x W x H: 2.4 x 2.4 x 2 m³). The temperature inside the cold room was maintained by a refrigeration machine (maximal compressor power 2.2 kW, refrigerant R404a). The cold room, its evaporator and instrumented sensors were shown in Fig.2. Cold supply air produced by the refrigeration machine was discharged at the evaporator outlet and circulated around the cold room before returning at the evaporator suction grid. The return air temperature, as a result of the heat exchange between the supply air, the product and the surrounding air inside the room, can be considered as the representative air temperature inside the cold room. The room was installed inside another climatic room, called “external cell”, with larger dimension, in order to simulate outdoor conditions. The external cell

temperature can be set between 17 °C (mild climate) and 30 °C (hot climate); these values were imposed by the technical limit related to the temperature control of this cell. Two types of sensor were used: temperature sensors (T-type thermocouples, uncertainty $\pm 0.1^\circ\text{C}$) and wattmeter (Digiwatt, precision 2 %). 17 thermocouples were installed at different positions (Fig. 2a): near the evaporator outlet (1 sensor, supply air temperature T_s), inside the product (8 sensors, product temperatures $T_{p,k}$, $k=1..8$), near the return air grid (3 sensors, return air temperatures $T_{r,j}$, $j=1..3$) and at different positions in the external cell (5 sensors, external air temperatures $T_{ext,i}$, $i=1..5$). The wattmeter was used to measure the electrical power demand P of the compressor. The measurements were performed with a regular period of 5s or 20s.

Product:

Test-packages (methylcellulose, 50 mm H; 100 mm W; 100 mm D) which are currently used for testing refrigerated cabinets according to ISO-23953-2 were used to simulate the thermal behavior of a real product. They were loaded in 8 boxes (Fig. 2b). Each box contained 32 test-packages and weighed 24.5 kg. In each box, one thermocouple was placed at the top surface of the product. This position was chosen because it is the position that is the most impacted in each box by the temperature variation inside the cold room.

Measurement protocol

Each experiment was composed of two main phases, namely DR period and stabilization (or restabilization) phases before the next DR period, as shown in Fig. 1. In the stabilization stage, the refrigeration unit was working; the room temperature was controlled by a compressor on/off regulation with thermostat sensor. During DR stage, the compressor was turned off while the fans of the refrigeration unit were maintained to ensure uniformed storage conditions [39]. DR was applied between two defrosting periods. Defrost is executed by hot gas. Automatic defrosting cycles occurred nearly every 6 h. After DR application, the compressor restarts and the temperature decreases and the condition before DR was reestablished. Several DR applications can be performed in one experiment.

4.2. Temperature evolution inside cold room – temporal features

Fig.1 shows the return air temperature evolution inside the cold room obtained by $T_{r,1}$, one of the 3 sensors placed near the return air grille (cf. Fig.2a). Three temporal features can be observed:

- (1) Regular fluctuations: the ambient temperature (-16.5 ± 1.1 °C) is controlled by the thermostat sensor and fluctuates around the set point temperature. These fluctuations are due to on/off compressor working cycles. The compressor working cycle's duration is not constant (340 ± 20 s). These discrepancies are related to the presence (or not) of frost in the evaporator.
- (2) DR period (3600 s in this example) with an important temperature rise due to the turning off of the compressor
- (3) Defrost period (duration ~ 600 s) with a smaller temperature rise.

Two time-dependent parameters related to DR and defrost periods were generated and used by the models:

$\delta_{t,DR}(t)$: time elapsed since the last DR period (see Fig.1)

$\delta_{t,def}(t)$: time elapsed since the last defrost period (see Fig.1)

These parameters will be used as reinforcement learning parameters in the models. They allow a better prediction of the cold room behaviour in the moments following the DR and defrost, when the condition of the cold room is not yet stabilized.

4.3. Data collection

Three experimental datasets, EP, E1 and E2 were used in this study. Fig. 3a presents the EP dataset: return air temperature T_r and product temperatures of the bottom ($T_{p,1}$) and top ($T_{p,4}$) boxes (cf. Fig.2). Three DR periods of respectively 1 h, 2 h and 3 h can be observed at time around 0.2 day, 1.1 day and 2 day. The longer the DR period, the higher is the air temperature peak: -7 °C for 1 h, -3 °C for 2 h and 0 °C for 3 h. The DR impact on product is less important than on air. Moreover, stronger impact was observed for the product at the top box ($T_{p,4}$) which attained -10 °C during the DR of 3 h. The product at the bottom ($T_{p,1}$) attained only -15 °C. The product initial temperatures were around -16 °C.

Because the DR impacts were much stronger on air than on products, only air temperatures were considered as inputs and outputs for the 2 datasets E1 (Fig. 3b) and E2 (Fig. 3c). Many peaks can be observed: DR periods and defrost periods (def, smaller peaks). The use of these two datasets allows different amounts of data, DR pattern and occurrence to be tested.

- Dataset E1 (210 300 points: 126 180 points for training and 84 120 points for test): larger dataset of 12 days with a measure interval of 5s, DR periods were imposed in a random manner, both in terms of duration (30 mins, 1 h, 2 h or 3 h) and occurrence (1 or 2 DR per day).
- Dataset E2 (17 840 points: 10 704 points for training and 7 136 points for test): smaller dataset of 4 days with a measure interval of 20 s, DR periods were imposed in a regular manner (same duration of 1 h, 3 h between 2 DR) and more frequent (6 DR per day).

The characteristics of the datasets are presented in Table 1.

Each dataset was divided into 2 phases: training (in blue in Fig.3, 60% of the data) and test (in green, 40% of the data); this repartition was done according to Simmhan and Noor [40]. In the training phase, both input and output data (cf. 5.1.2. Configurations of inputs and outputs) were used to build the model while in the test phase; only input data were used for the prediction.

5. LSTM deep learning models development

5.1. Formalization of the system behaviour prediction problem with time series

5.1.1. Prediction problem

The prediction problem of the cold room behaviour and its temporal features, during DR periods in particular, is represented by Eq. 1:

$$y_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-m+1}, x_t, x_{t-1}, \dots, x_{t-p+1}) \quad (\text{Eq.1})$$

with y_t and x_t time series observations representing respectively output and input parameters, f a non-linear prediction function with $m + p$ inputs.

This prediction problem is equivalent to find the optimal \hat{f} which minimizes:

$$\text{error} = (\hat{y} - y) \quad (\text{Eq.2})$$

with $\hat{y}(m, H)$ an estimation of y for a given horizon H from a series of m observations $y_t, y_{t-1}, \dots, y_{t-m+1}$.

5.1.2. Configurations of inputs and outputs

The inputs x_t (multiple time series) consist of: air temperature in the external cell T_{ext} (°C); supply air temperature T_s (°C) (optional input); $\delta_{t,DR}$ (s) and $\delta_{t,def}$ (s): time elapsed since the last DR and defrost periods respectively; and three binary inputs Comp, Def and DR which indicate the state (ON or OFF) of the compressor, defrost and DR control at the moment t . These data were measurements (T_{ext} and T_s) or processed data from measurements (other inputs).

The outputs y_t (multiple time series) are the return air temperature T_r (°C), the compressor power demand P (W) and product temperature T_p (°C) (optional output).

The list of input and output was presented in Table 2a.

In order to determine the influence of the type and number of inputs and outputs on the quality of prediction, 4 configurations were tested (Table 2b). These configurations have 5 common inputs ($\delta_{t,DR}$, $\delta_{t,def}$, Comp, Def and DR) and 1 common output (P). The difference lies in the T_{ext} (input) and T_r (output). The configurations P and 1 use the N_e time series of N_e temperature sensors ($T_{ext,i}$, $i=1..N_e$) that measure the air temperature of the external cell at various positions as input while configurations 2 and 3 use $\overline{T_{ext}}$, the average value of $T_{ext,i}$.

$$\overline{T_{ext}} = \frac{1}{N_e} \sum_{i=1}^{N_e} T_{ext,i} \quad (\text{Eq.3})$$

In the same manner, the configurations P and 1 have as outputs N_r time series of the return air temperature ($T_{r,j}$, $j=1..N_r$) corresponding to the N_r sensors that measure this temperature while the configurations 2 and 3 predict the average return air temperature $\overline{T_r}$.

$$\overline{T_r} = \frac{1}{N_r} \sum_{j=1}^{N_r} T_{r,j} \quad (\text{Eq.4})$$

The product temperatures were considered as additional outputs in the configuration P. Only for the configuration 3, the supply air temperature T_s was used as additional input (Table 2b). In this study, $N_e=5$ and $N_r=3$.

5.2. LSTM models development

4 LSTM models: traditional LSTM, Convolutional LSTM, Stacked LSTM and Bidirectional LSTM were developed using the Keras deep learning library [41].

5.2.1. Traditional LSTM

The traditional LSTM model has been used mostly for time series study; it allows both short and long terms information to be kept by using specialized gating and memory mechanism. An LSTM model is composed principally of LSTM units (or cells) to store and process information. Fig.4a presents the elements of a unit. The relations between those elements are described by following equations:

$$f_t = \sigma_A(W_f x_t + U_f h_{t-1} + b_f) \quad (\text{Eq.5})$$

$$i_t = \sigma_A(W_i x_t + U_i h_{t-1} + b_i) \quad (\text{Eq.6})$$

$$o_t = \sigma_A(W_o x_t + U_o h_{t-1} + b_o) \quad (\text{Eq.7})$$

$$\tilde{c}_t = \sigma_A(W_c x_t + U_c h_{t-1} + b_c) \quad (\text{Eq.8})$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (\text{Eq.9})$$

$$h_t = o_t \circ \sigma_A(c_t) \quad (\text{Eq.10})$$

with x_t : input vector, h_t : hidden state (output) vector, f_t : forget gate's activation vector, i_t : input gate's activation vector, o_t : output gate's activation vector, \tilde{c}_t : cell input activation vector, c_t : cell state vector, W and U: weight matrices, b: bias vector, σ_A : activation function, operator \circ indicates the Hadamard (element-wise) product.

A LSTM layer is formed by connecting a number of LSTM units (Fig. 4b). As shown in Fig. 5a, the LSTM model is composed of a LSTM layer, a dropout layer and dense layers. The dropout regularization technique is used to avoid overtraining by randomly selecting ignored neurons (dropped-out) during training. Dense layer is a regular layer of neurons; each neuron receives input from all the neurons in the previous layer, thus densely connected its input to the output. More information on these layers can be found in the Keras website [41].

After an important number of tests, the following parameters were selected for the LSTM model:

- Number of units, 1024 for configuration 1 and 512 for configurations 2 and 3 (since less data were used as inputs for these configurations, less units were needed). The choice of these numbers is done as a compromise between the learning time and the quality of obtained prediction.
- A linear (Scaled Exponential Linear Unit - SeLu) activation function

$$\sigma_A(x) = \lambda \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (\text{Eq.11})$$

With $\lambda = 1.0507$ and $\alpha = 1.67326$.

- A cost function using the Root Mean Square Error (RMSE) to measure the difference between the predicted value and the actual value during learning in order to optimize weight parameters.

5.2.2. Stacked LSTM

The Stacked LSTM model was formed by using several LSTM layers. It was developed in order to answer the following question: could a better prediction be reached by increasing the depth of the network? After several tests, for the configuration 1, the selected model consists of a stack of three layers of LSTM, namely: a first layer of 1024 memory units, a second layer of 512 memory units and a third layer of 256 memory units. For the configurations 2 and 3, the model is formed by 2 LSTM layers: a first layer of 512 memory units and a second layer of 256 memory units. Each layer of the network is separated from the next layer by a dropout layer, allowing less overtraining and robust generalization results. Each LSTM layer used the same parameters as the traditional LSTM model. The structure of the stacked LSTM model for the configuration 2 (or 3) was shown in Fig. 5b.

5.2.3. Bidirectional LSTM

This kind of model is an extension of traditional LSTM that can improve the model performance by linking past and future data. The bidirectional LSTM model is composed of 2 Bi-LSTM layers (512 units for the 3 configurations) and dense layers (Fig. 5c). The Bi-LSTM layer trains two instead of one LSTM layer on the input sequence. The first on the input sequence related to past data and the second on a reversed copy of the input sequence (backward-future data).

5.2.4. Convolutional LSTM

The convolutional LSTM network is the combination of CNN and LSTM networks. It was developed with the following parameters: 40 filters, corresponding to the outputs of the convolutional part of the model; a kernel size of 2x10, corresponding to the dimensions of the convolution window; a normalization layer, allowing the activations of the convolutional LSTM layer to be normalized. Fig. 5d shows the structure of this model which is the same for the 3 configurations.

5.3. Error estimation for model performance assessment

In order to compare the performance of the various models and simulation configurations, Fit and MAE criteria were implemented and computed over the prediction horizon $[t + 1; t + H]$ in which there are m observations.

Fit criterion measures the distance between the predicted values \hat{Y} and reference (observations) Y (\bar{Y} is the average of Y).

$$\text{Fit}(\%) = 100 \times \left(1 - \frac{|\hat{Y} - Y|}{|Y - \bar{Y}|} \right) \quad (\text{Eq.12})$$

The closer Fit is to 100%, the more it indicates a correctly predicted variable.

The Mean Absolute Error (MAE) is the arithmetic mean of the differences between predicted and actual observations.

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |\hat{Y}_i - Y_i| \quad (\text{Eq.13})$$

Unlike Fit, a smaller MAE corresponds to a better predicted variable.

6. Results and discussion

6.1. Prediction performance of 4 LSTM models in training and test phases

6.1.1. Air temperature and power demand

The comparison between reference (experimental) data and numerical prediction around a DR period in training phase obtained by traditional LSTM model is shown in Fig.6; the dataset E1 and the configuration 1 (Cf. Table 2b) was used for the simulations. The evolution of temperature and power demand was well predicted, even during and after DR period when sudden changes occurred. Indeed, during the DR period of 1 h, a peak of air temperature rise (from -15°C to -6°C) was observed and the compressor was stopped (compressor power demand = 0). Just after DR period, the temperature decreased to attain its level before DR while the compressor had to work harder (the first compressor on-off cycle after DR, at time around 2 h, was longer than regular cycles). Prediction noises were observed during DR period in both temperature and power curves. They might correspond to the temperature and power fluctuations during regular cycles (temporal feature number 1 in Fig. 1).

Fig.7 presents the comparison between experimental data and numerical prediction around a DR period in test phase obtained by the 4 LSTM models for the return air temperature. The prediction performance was lower than during training phase with stronger noises. The sharp temperature peaks which can be observed at the beginning of DR period (time = 1h) in Traditional and Stacked LSTM prediction curves might be related to defrost periods which were learned by the models during training phase. Higher performance was obtained by the 4 models for the prediction of compressor demand power during test phase (Fig.8) in which there are less differences between training and test phases. For the configuration 1 and dataset E1, more noises are observed in the prediction curves of Bidirectional and Convolutional LSTM (Fig.7 and 8) during DR period than for Traditional and Stacked LSTM.

The prediction performances of the 4 LSTM models (for the configuration 1 and dataset E1) during training and test phases are shown in Table 3a (Fit) and Table 3b (MAE). While it is obvious that the test phase has lower performance than the training phase, as previously observed, temperature prediction shows lower performance than power demand prediction for both phases. The reason is that temperature variation is associated with more complex features (Cf. 4.2. Temperature evolution inside cold room – temporal features) and thus it is more difficult for the models to predict the temperature behaviour than the power demand behaviour. While the best Fit (closest to 100 %) during training phase is obtained with stacked LSTM (79.85% for return air temperature T_r and 93.08% for compressor power demand P), it's traditional LSTM that has the best Fit during test phase (55.06% for T_r and 91.22% for P). Table 3b shows that the best MAE values (smallest ones) are obtained by traditional LSTM for both training and test phases. In general, Bidirectional and Convolutional LSTM show lower performances (both Fit and MAE) for these simulations.

6.1.2. Product temperature

The comparison between experimental and numerical prediction by 4 models for product temperatures in test phase is presented in Fig. 9. The dataset EP (cf. table 1) and the configuration P (cf. Table 2b) were used. As shown in Fig. 3a and section 4.3 (Data collection), this dataset has 3 DR periods of respectively 1 h, 2 h and 3 h. The 2 first DR periods of 1 h and 2 h were used in training while the last DR period of 3 h were used in test phase. Two product temperatures were shown: $T_{p,1}$ from the bottom box and $T_{p,4}$ from the top box. Two different behaviours were observed in the experimental data (finer lines): higher temperature rise during DR for products inside the top box ($T_{p,4}$) which were more impacted by the air temperature rise inside the cold room than products inside the bottom box ($T_{p,1}$). While the 4 models were able to predict this difference between product temperatures at bottom and top boxes, it is noted that a better performance was obtained by convolutional LSTM: only this model predicted the temperature peak at the end of the DR period of 3 h. For the 3 other LSTM models (traditional, stacked and bidirectional), the prediction was good during the first 2 h but these

models have failed to predict that the product temperature continued to rise up after. Two explanations can be given. First, convolutional LSTM is more proficient than other LSTM models in extracting complex features. Second, as only 2 DR periods of 1 h and 2 h were used in training, other models have not yet learned the behaviour of longer DR period so that their prediction was only good for the first 2 h. In the next sections, 2 datasets of longer time with more frequent DR periods will be investigated. One of the objectives is to see which models are more sensitive to the variations of temperature. As DR impacts are more visible on air temperature than on product temperatures, only return air temperature T_r will be considered in these analysis.

6.2. Influence of data availability on model prediction performance

Fig. 10 presents the results obtained by traditional LSTM model in test phase for three configurations of inputs and outputs 1, 2 and 3 (cf. Table 2b). Compared to the configuration 1 that used various time series from different sensors ($T_{ext,i}$, $i=1..5$: 5 time series and $T_{r,j}$, $j=1..3$: 3 time series), the second configuration used the average value $\overline{T_{ext}}$ (1 time series) and $\overline{T_r}$ (1 time series) and thus, less information were available for model building. As a consequence, the results obtained by the configuration 2 are of lower quality than for the configuration 1, in particular for the temperature prediction in which the configuration 2 predicted a much higher temperature rise during DR. However, due to the use of average values, lower noise is observed in the curves of configuration 2 compared to the curves of configuration 1.

The configuration 3 used the same inputs and outputs as in the configuration 2 with an additional input: the supply air temperature T_s (1 time series). The results of this configuration outperformed those obtained by both configurations 1 and 2. It is emphasized that from the point of view of thermal analysis, the supply air temperature contains already important part of the prediction as it brings rich information to the model in terms of air temperature and refrigeration machine behaviour. However, in this study, this data was used only in this configuration in order to determine to what extent it can improve the prediction.

The prediction performances (Fit and MAE) of the 3 configurations by the 4 LSTM models are shown in Table 4a and b for the test phase. For the 4 models, highest performances (highest Fit and lowest MAE) are obtained for the configuration 3 while lowest performances are observed for the configuration 2. This result shows the importance of data availability: more information from more sensors (configuration 1) or from an optional sensor (configuration 3) leads to better predictions. The performances of the 4 models are quite similar for both Fit and MAE criteria. Nevertheless, for configurations 2 and 3, Bidirectional and Convolutional LSTM show higher performance than Stacked LSTM which is not the case for configuration 1. This might be explained by the fact that Bidirectional and Convolutional LSTM models are more sensitive to noise and as the configurations 2 and 3 use average values, these two models are less exposed to noise so that they can attain better performances.

6.3. Influence of the data characteristic (quality, quantity and DR pattern) on model performance

In order to identify the influence of the data characteristic on the performance of the 4 models, the results from a new dataset will be analyzed. This dataset E2 (17 840 points) is 10 times smaller than the dataset E1 (210 300 points). However, as shown in Table 5 and Fig. 11, the prediction performances are of the same order of magnitude as those from the dataset E1 (Cf. Table 4). For the configuration 2 in particular, even higher performances (higher Fit and lower MAE) were obtained. In Fig. 11, it can be observed that the temperature rise during DR was better predicted for the configuration 2 than the one with the dataset E1 in Fig. 10. These good performances can be explained

by the DR pattern in this dataset which is much more regular and more frequent than in the dataset E1. For the dataset E2, the performances of the 4 models are also similar: traditional LSTM and stacked LSTM gave better performance for the configuration 1 while higher performances were obtained by bidirectional and convolutional LSTM for the configurations 2 and 3.

7. Conclusion

This study is one of the first works on demand response (DR) in cold rooms and warehouses. Its aim is to develop “data driven” models for the application of DR. Many questions have been addressed such as which models can be applied as ANNs are a very large family with many branches? How many data are needed for training? Are the models sensible to data quality (noise)? Four deep learning models, traditional LSTM, stacked LSTM, bidirectional LSTM and convolutional LSTM, were developed to predict temperature and power demand variations for the application of demand response (DR, by stopping the compressor of the refrigeration machine) in a cold room. Model building data were collected from experimentation in a laboratory cold room. Four configurations of inputs and outputs from different combinations of sensors were applied to identify the impact of data availability on the prediction. It was found that while using more sensors leads to better predictions, it is more important to place the sensor at a suitable position that can give rich information for model learning. In order to study the impact of data characteristic on the prediction performance, two different datasets were used: one with high number of points but random DR occurrence and one with fewer points (10 times less) but more frequent and regular DR occurrence. Similar performance was obtained by two datasets which shows that if the temporal features of the time series are regular, it is not necessary to have a large dataset. However, DR occurrence might be random in industrial conditions which demands important quantity of data to cover various features related to temperature and power demand variations. The 4 developed deep learning LSTM models showed similar performances: all of the developed models could be used for DR prediction problem. Convolutional LSTM model has the best performance in predicting product temperatures. Bidirectional LSTM and convolutional LSTM models are more sensitive to noise so that their performances are better when data with less noise are used. The results of this work show that in practice, ANNs models, LSTM family in particular, can be used to predict product and air temperature evolutions inside cold room, even during DR period when different behaviours occur. The present study is a preliminary one, to be followed by a second step towards the identification of a more proper set of variables that influence the cold chain behaviour under DR protocols. Moreover, the comparison of the advantages and drawbacks of physical model (ex: conventional lumped thermal capacity model) and data driven models (ANNs) will be the objective of our further works. Finally, the developed models using data obtained in a laboratory cold room will be extended for application in cold warehouses.

8. Acknowledgements

This work was funded by the Agency for the Environment and Management (ADEME) [Project ADEME n°1581C0116 – Flexifroid - Effacement des entrepôts frigorifiques de surgelés : évaluation de l’impact énergétique et du risque produit, 2016 - 2019].

References

[1] Renewable Energy Directive (RED II), Directive (EU) 2018/2001 (recast) on the promotion of the use of energy from renewable sources.

- [2] M. McPherson, B. Stoll. Demand response for variable renewable energy integration: A proposed approach and its impacts. *Energy*. 197 (2020) 117205.
- [3] M. Glavan, D. Gradišar, S. Moscariello, Đ. Juričić, D. Vrančić. Demand-side improvement of short-term load forecasting using a proactive load management – a supermarket use case. *Energy and Buildings*. 186 (2019) 186-94.
- [4] N. O'Connell, P. Pinson, H. Madsen, M. O'Malley. Benefits and challenges of electrical demand response: A critical review. *Renewable and Sustainable Energy Reviews*. 39 (2014) 686-99.
- [5] IIR. Cold chain technology brief: cold storage and refrigerated warehouse. International Institute of Refrigeration 2018.
- [6] A. Ambaw, N. Bessemans, W. Gruyters, S.G. Gwanpua, A. Schenk, A. De Roeck, M.A. Delele, P. Verboven, B.M. Nicolai. Analysis of the spatiotemporal temperature fluctuations inside an apple cool store in response to energy use concerns. *International Journal of Refrigeration*. 66 (2016) 156-68.
- [7] A.R. East, N.J. Smale, F.J. Trujillo. Potential for energy cost savings by utilising alternative temperature control strategies for controlled atmosphere stored apples. *International Journal of Refrigeration*. 36 (2013) 1109-17.
- [8] M. Akerma, H.-M. Hoang, D. Leducq, A. Delahaye. Experimental characterization of demand response in a refrigerated cold room. *International Journal of Refrigeration*. 113 (2020) 256-65.
- [9] E. Dermesonluoglu, G. Katsaros, M. Tsevdou, M. Giannakourou, P. Taoukis. Kinetic study of quality indices and shelf life modelling of frozen spinach under dynamic conditions of the cold chain. *Journal of Food Engineering*. 148 (2015) 13-23.
- [10] Y. Phimolsiripol, U. Siripatrawan, D.J. Cleland. Weight loss of frozen bread dough under isothermal and fluctuating temperature storage conditions. *Journal of Food Engineering*. 106 (2011) 134-43.
- [11] V. Vicent, F.T. Ndoye, P. Verboven, B.M. Nicolai, G. Alvarez. Quality changes kinetics of apple tissue during frozen storage with temperature fluctuations. *International Journal of Refrigeration*. 92 (2018) 165-75.
- [12] H.-M. Hoang, S. Duret, D. Flick, O. Laguerre. Preliminary study of airflow and heat transfer in a cold room filled with apple pallets: Comparison between two modelling approaches and experimental results. *Applied Thermal Engineering*. 76 (2015) 367-81.
- [13] F.J. Montáns, F. Chinesta, R. Gómez-Bombarelli, J.N. Kutz. Data-driven modeling and learning in science and engineering. *Comptes Rendus Mécanique*. 347 (2019) 845-55.
- [14] H. Fu, W.-J. Cheng, Y. Qin. Exploration of data-driven methods for multiphysics electromagnetic partial differential equations. *IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)2020*.
- [15] B. Maschler, M. Weyrich. Deep Transfer Learning for Industrial Automation: A Review and Discussion of New Techniques for Data-Driven Machine Learning. *IEEE Industrial Electronics Magazine* 2020.
- [16] A. Afram, F. Janabi-Sharifi. Review of modeling methods for HVAC systems. *Applied Thermal Engineering*. 67 (2014) 507-19.

- [17] M. Akerma, H.-M. Hoang, D. Leducq, P. Clain, A. Delahaye. Demand response in refrigerated warehouse. 4thIEEE International Smart Cities Conference, Kansas City, USA 2018.
- [18] J.E. Altwies, D.T. Reindl. Passive thermal energy storage in refrigerated warehouses. *International Journal of Refrigeration*. 25 (2002) 149-57.
- [19] S. Estrada-Flores, A.C. Cleland, D.J. Cleland. Prediction of the dynamic thermal behaviour of walls for refrigerated rooms using lumped and distributed parameter models. *International Journal of Refrigeration*. 24 (2001) 272-84.
- [20] T. Berthou, P. Stabat, R. Salvazet, D. Marchio. Development and validation of a gray box model to predict thermal behavior of occupied office buildings. *Energy and Buildings*. 74 (2014) 91-100.
- [21] M. Hu, F. Xiao, L. Wang. Investigation of demand response potentials of residential air conditioners in smart grids using grey-box room thermal model. *Applied Energy*. 207 (2017) 324-35.
- [22] S. Tian, S. Shao, B. Liu. Investigation on transient energy consumption of cold storages: Modeling and a case study. *Energy*. 180 (2019) 1-9.
- [23] S.S. Pappas, L. Ekonomou, D.C. Karamousantas, G.E. Chatzarakis, S.K. Katsikas, P. Liatsis. Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. *Energy*. 33 (2008) 1353-60.
- [24] P. Sen, M. Roy, P. Pal. Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy*. 116 (2016) 1031-8.
- [25] C. Yuan, S. Liu, Z. Fang. Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model. *Energy*. 100 (2016) 384-90.
- [26] W. Scherrer, M. Deistler. Chapter 6 - Vector autoregressive moving average models. in: H.D. Vinod, C.R. Rao, (Eds.), *Handbook of Statistics*. Elsevier2019. pp. 145-91.
- [27] T. Ahmad, H. Chen. Deep learning for multi-scale smart energy forecasting. *Energy*. 175 (2019) 98-112.
- [28] L. Yin, Q. Gao, L. Zhao, T. Wang. Expandable deep learning for real-time economic generation dispatch and control of three-state energies based future smart grids. *Energy*. 191 (2020) 116561.
- [29] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*. 50 (2009) 90-6.
- [30] S. Hochreiter, J. Schmidhuber. Long Short-Term Memory. *Neural Computation*. 9 (1997) 1735-80.
- [31] Y. LeCun, Y. Bengio, G. Hinton. Deep learning. *Nature*. 521 (2015) 436-44.
- [32] C. Xu, H. Chen, J. Wang, Y. Guo, Y. Yuan. Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method. *Building and Environment*. 148 (2019) 128-35.
- [33] T.-Y. Kim, S.-B. Cho. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*. 182 (2019) 72-81.

- [34] A. Das, M.K. Annaqeeb, E. Azar, V. Novakovic, M.B. Kjærgaard. Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Applied Energy*. 269 (2020) 115135.
- [35] R. Lu, S.H. Hong. Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Applied Energy*. 236 (2019) 937-49.
- [36] L. Wen, K. Zhou, S. Yang. Load demand forecasting of residential buildings using a deep learning model. *Electric Power Systems Research*. 179 (2020) 106073.
- [37] G. Onoufriou, R. Bickerton, S. Pearson, G. Leontidis. Nemesyst: A hybrid parallelism deep learning-based framework applied for internet of things enabled food retailing refrigeration systems. *Computers in Industry*. 113 (2019) 103133.
- [38] C. Zhang, H. Zhang, J. Qiao, D. Yuan, M. Zhang. Deep Transfer Learning for Intelligent Cellular Traffic Prediction Based on Cross-Domain Big Data. *IEEE Journal on Selected Areas in Communications*. 37 (2019) 1389-401.
- [39] R.W. Brooks, R.H. Green, A.J. Cleland, P. Senior. The potential of refrigerated stored for electricity load management 20th International Congress of Refrigeration, Sydney, 1999. pp. 1289-96
- [40] Y. Simmhan, M.U. Noor. Scalable Prediction of Energy Consumption using Incremental Time Series Clustering. *IEEE International Conference on Big Data2013*.
- [41] Keras. <https://keras.io/api/layers/>, last accessed on 03/07/2020

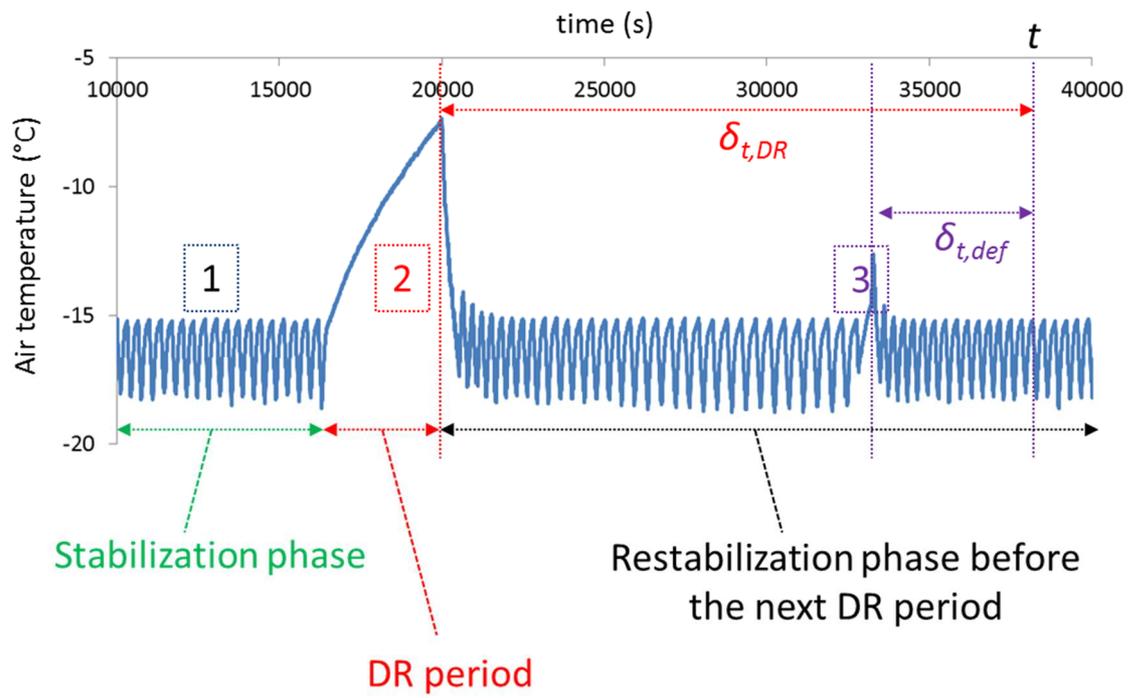


Figure 1: Return air temperature evolution and temporal features, 1 - compressor on-off regular cycles, 2 - DR period and 3 - defrost period (experimental data)

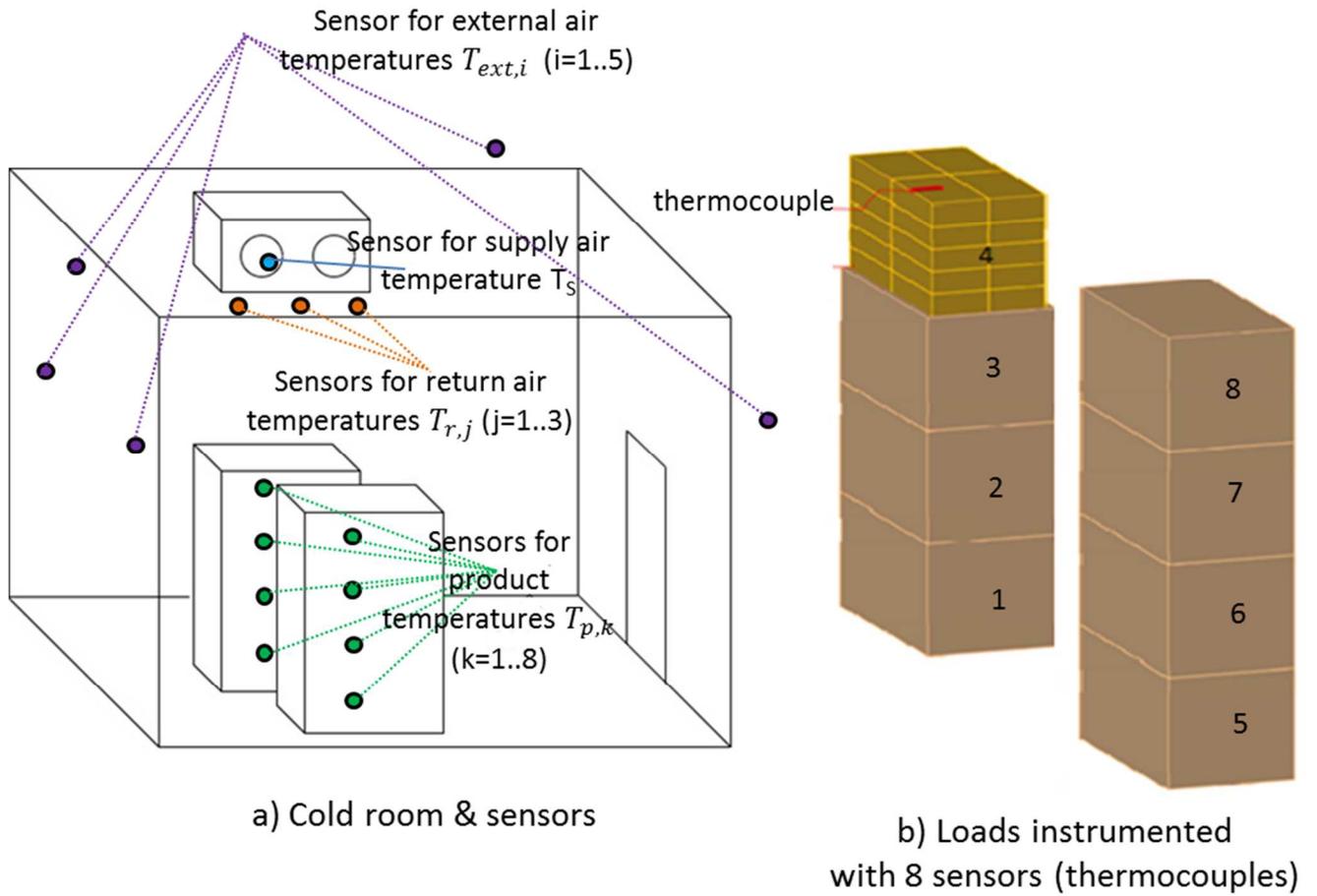
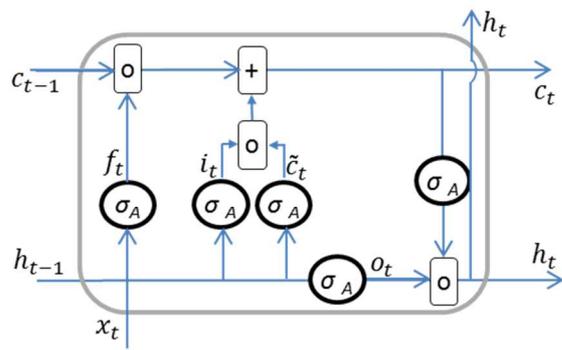
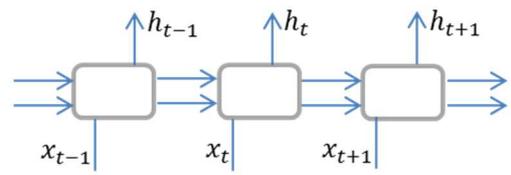


Figure 2: Cold room and temperature sensors



a. LSTM unit (cell)



b. LSTM layer

Figure 4: LSTM unit and layer

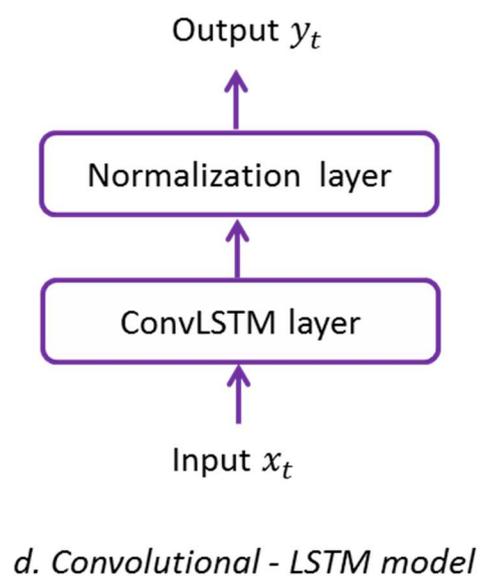
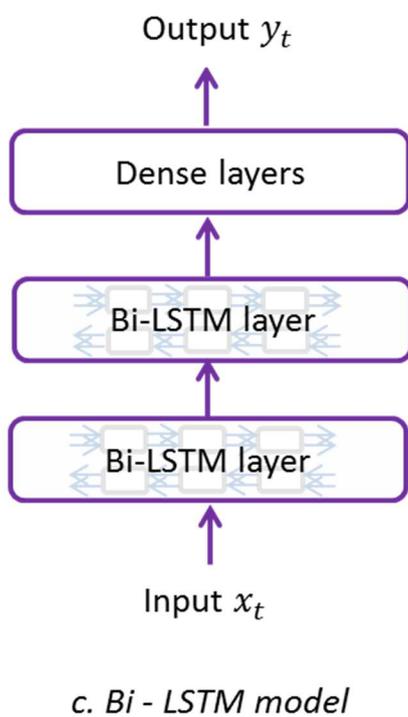
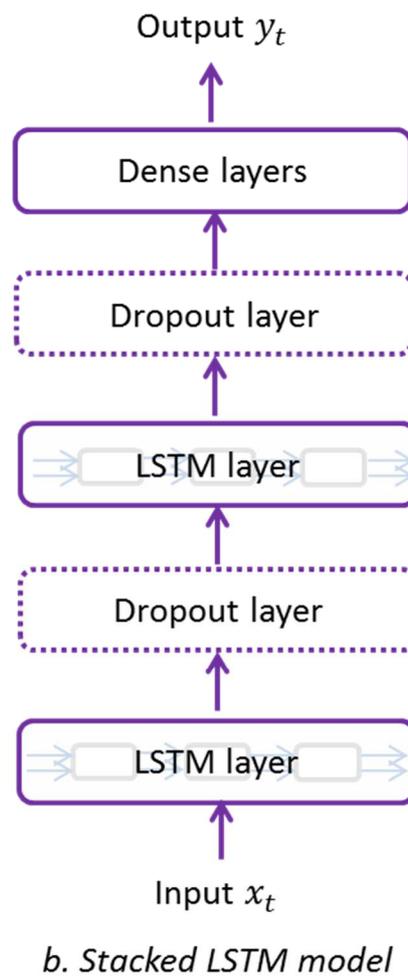
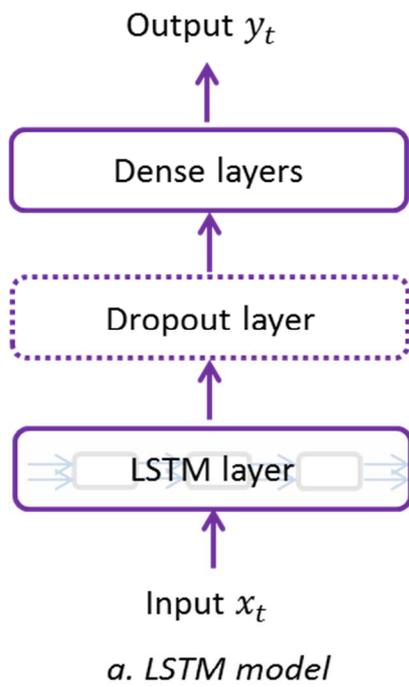


Figure 5: Developed LSTM models

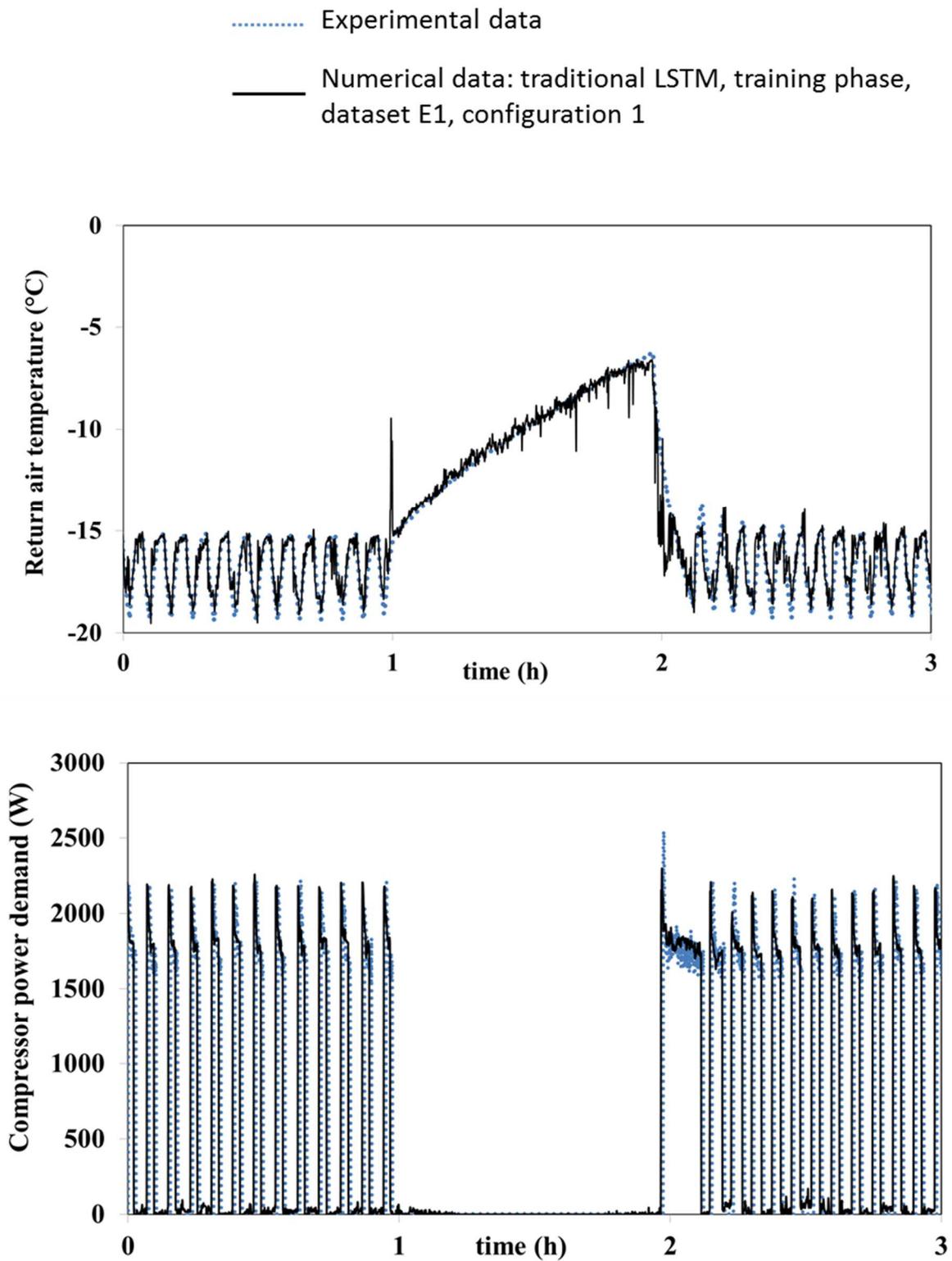


Figure 6: Comparison between experimental and numerical prediction before, during and after DR period in training phase obtained by traditional LSTM model using dataset E1 and configuration 1

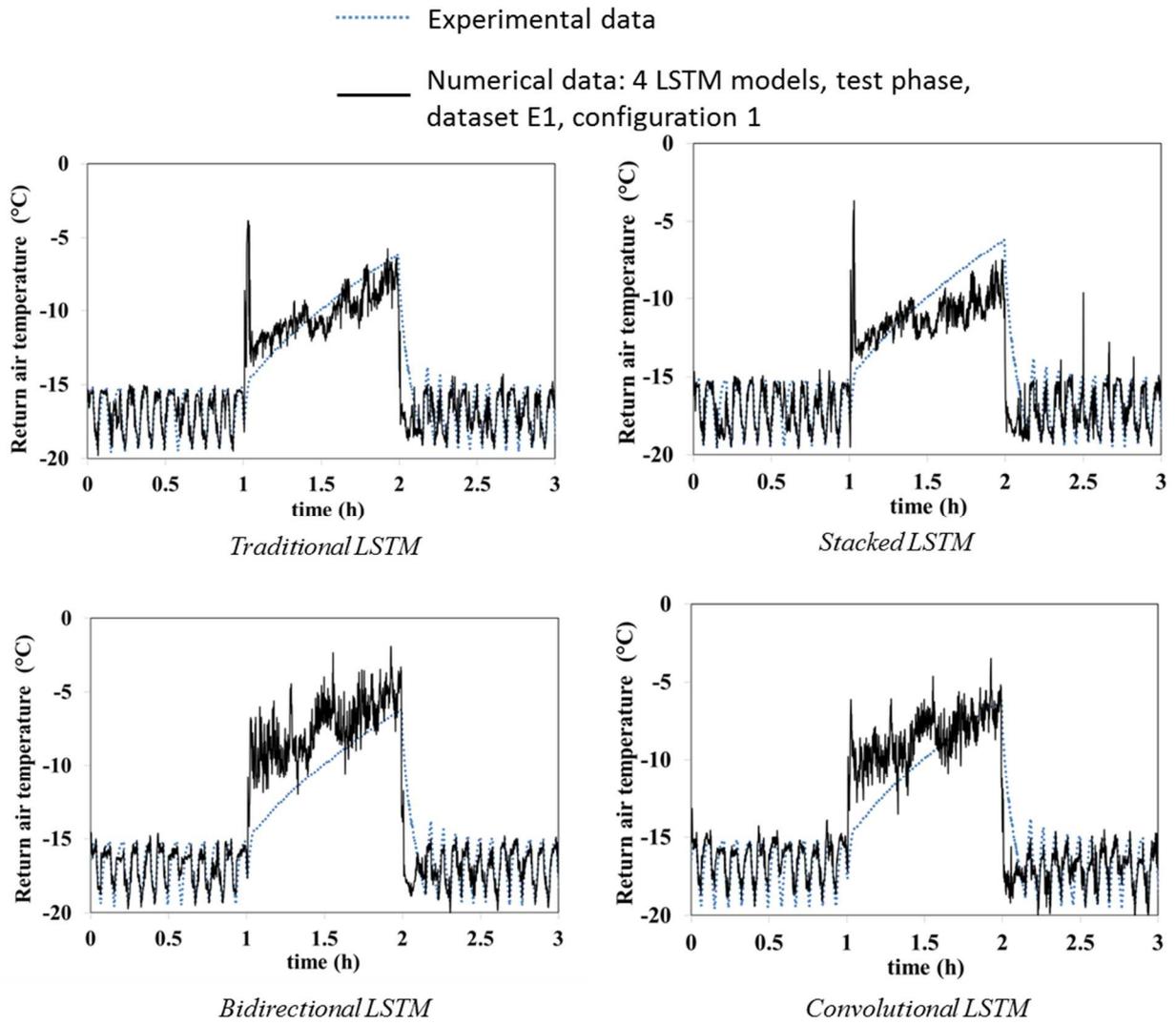


Figure 7: Comparison between experimental and numerical prediction for return air temperature before, during and after DR period in test phase obtained by 4 models (dataset E1 and configuration 1)

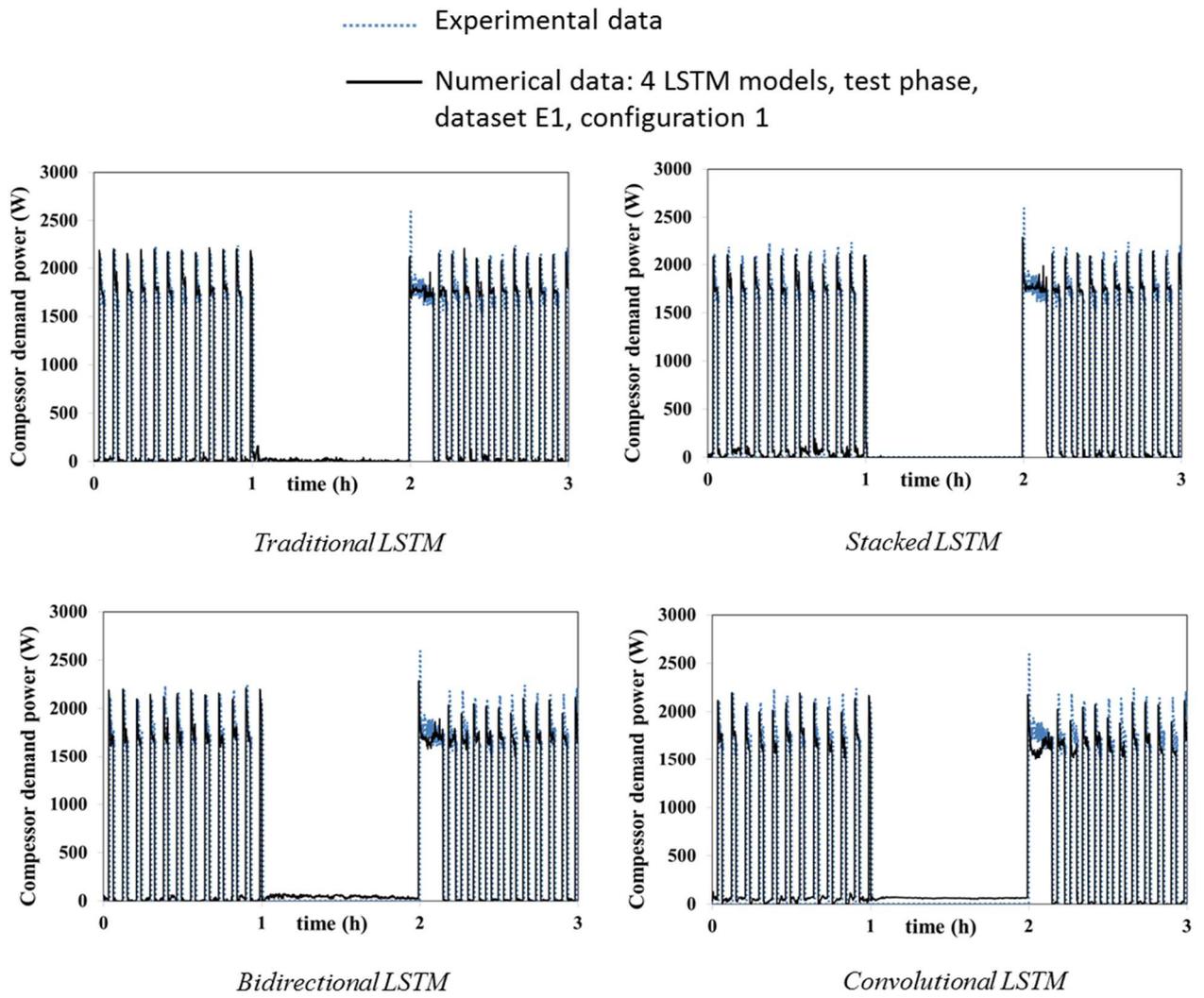
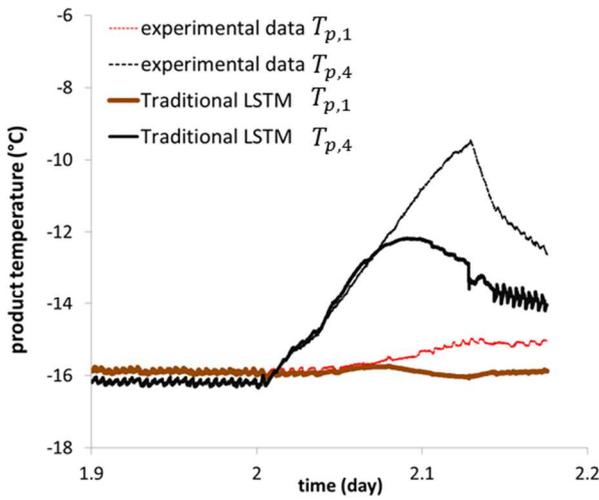
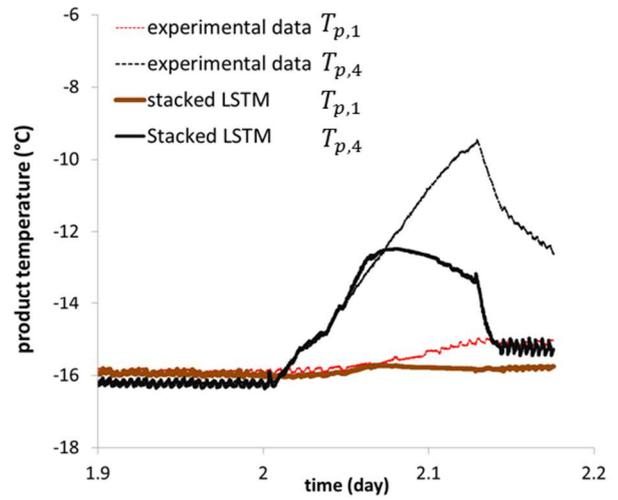


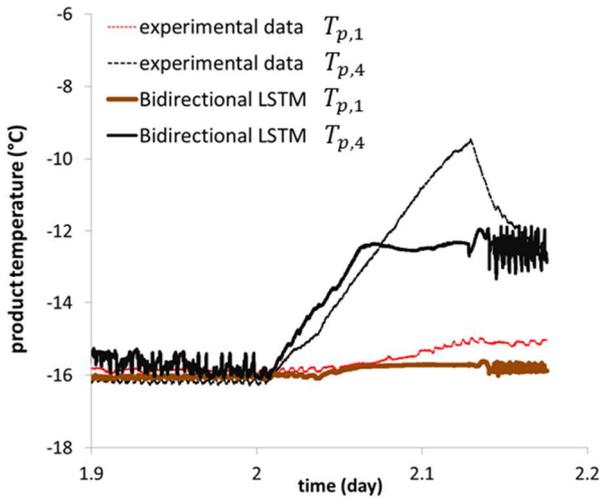
Figure 8: Comparison between experimental and numerical prediction for compressor demand power before, during and after DR period in test phase obtained by 4 models (dataset E1 and configuration 1)



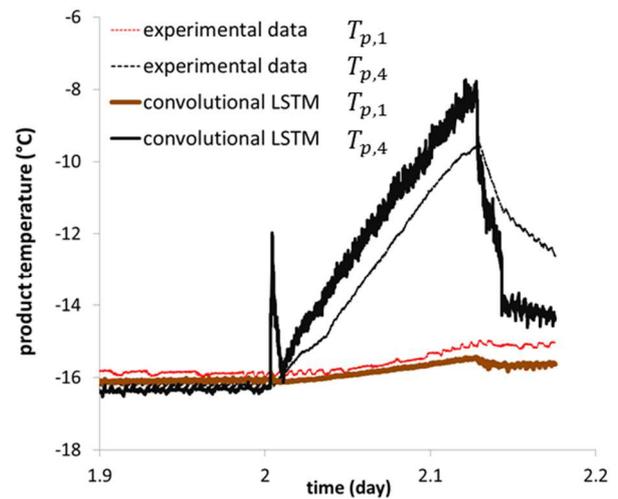
Traditional LSTM



Stacked LSTM



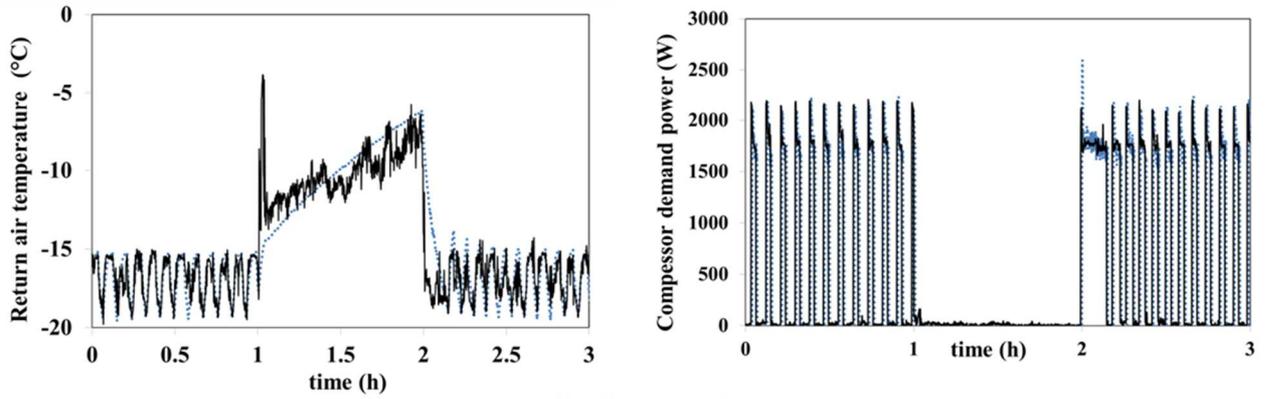
Bidirectional LSTM



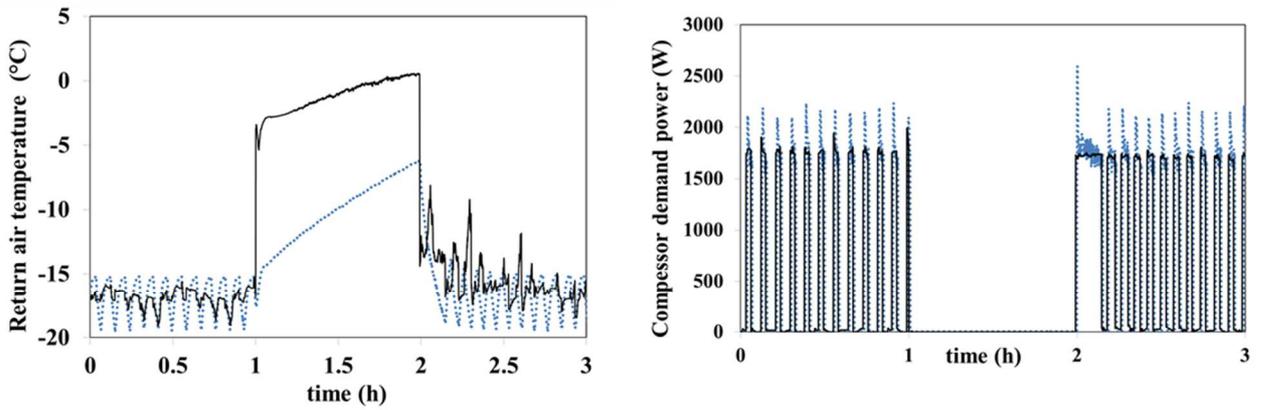
Convolutional LSTM

Figure 9: Comparison between experimental and numerical prediction for product temperatures in test phase obtained by 4 models (dataset EP and configuration P)

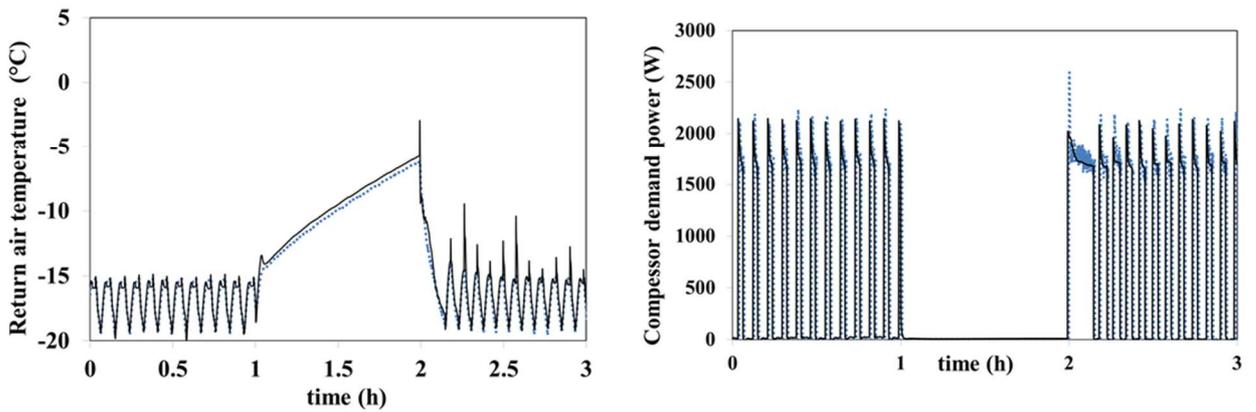
----- Experimental data
——— Numerical data: traditional LSTM, test phase, dataset E1, 3 configurations



Configuration 1



Configuration 2



Configuration 3

Figure 10: Comparison between experimental and numerical prediction before, during and after DR period in test phase obtained by traditional LSTM model using dataset E1 and 3 configurations

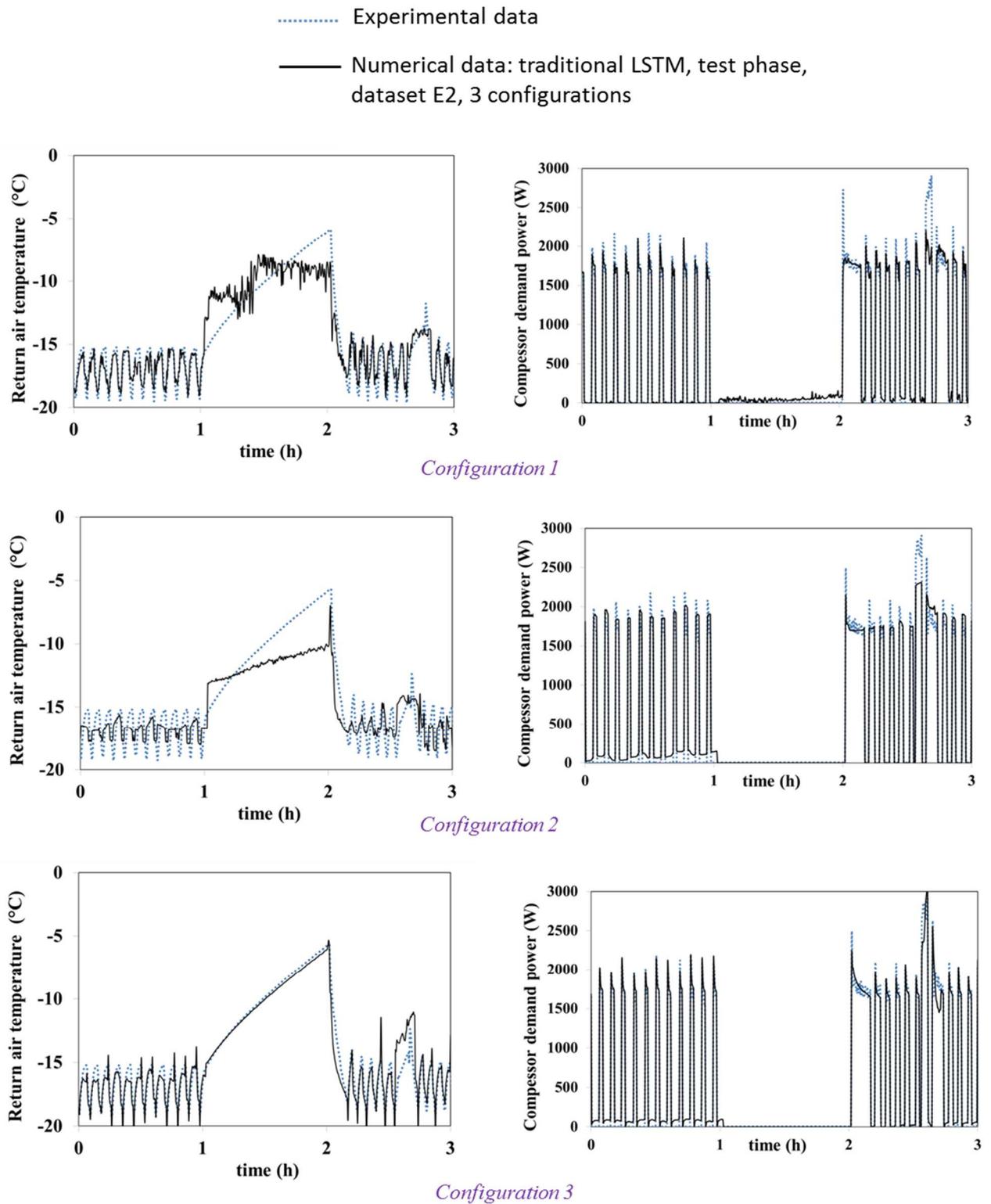


Figure 11: Comparison between experimental and numerical prediction before, during and after DR period in test phase obtained by traditional LSTM model using dataset E2 and 3 configurations

	Dataset EP	Dataset E1	Dataset E2
Number of points for training	25 600	126 180	10 704
Number of points for test	12 000	84 120	7 136
Total number of points	37 600	210 300	17 840
DR pattern	regular	random	regular
DR duration	1 h, 2 h, 3 h	30 mins, 1 h, 2 h, 3 h	1 h
DR occurrence	1 per day	1 or 2 DR per day	6 DR per day
Measure interval	5 s	5 s	20 s
Measurement duration	2.2 days	12 days	4 days

Table 1: Data characteristics of three experimental datasets

	Name	Description
Input (x_i)	T_{ext}	External air temperature (in external cell), °C
	T_S (optional)	Supply air temperature, °C
	$\delta_{t,def}$	time elapsed since the last defrost period, s
	$\delta_{t,DR}$	time elapsed since the last DR period, s
	Comp	Comp=0 : compressor OFF ; Comp=1 : compressor ON, binary
	Def	Def=0 : Defrost OFF ; Def=1 : Defrost ON, binary
	DR	DR=0 : not in DR period; DR=1 : in DR period, binary
Output (y_i)	T_r (°C)	Return air temperature, °C
	P (W)	Compressor power demand, W
	T_p (optional)	Product temperature, °C

a. List of possible inputs and outputs

	Configuration P	Configuration 1	Configuration 2	Configuration 3
Input (x_i)	$T_{ext,i}$ (i=1..5)	$T_{ext,i}$ (i=1..5)	$\overline{T_{ext}}$ (average)	$\overline{T_{ext}}$ (average)
	$\delta_{t,def}$	$\delta_{t,def}$	$\delta_{t,def}$	$\delta_{t,def}$
	$\delta_{t,DR}$	$\delta_{t,DR}$	$\delta_{t,DR}$	$\delta_{t,DR}$
	Comp	Comp	Comp	Comp
	Def	Def	Def	Def
	DR	DR	DR	DR
Output (y_i)				T_S
	$T_{r,j}$ (j=1..3)	$T_{r,j}$ (j=1..3)	$\overline{T_r}$ (average)	$\overline{T_r}$ (average)
	P	P	P	P
	$T_{p,k}$ (k=1..8)			

b. Four configurations of inputs and outputs

Table 2: List of inputs and outputs and the configurations used in model development

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature $T_{r,1}$	Training	77.64	79.85	69.99	65.59
	Test	55.06	53.58	51.93	52.13
Compressor demand power P	Training	92.62	93.08	92.31	91.39
	Test	91.22	85.35	67.05	88.25

a. Fit (%)

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature $T_{r,1}$	Training	0.153	0.163	0.210	0.299
	Test	0.230	0.248	0.276	0.277
Compressor demand power P	Training	0.052	0.053	0.054	0.059
	Test	0.052	0.073	0.083	0.065

b. MAE

Table 3: Prediction performance (Fit and MAE) of 4 LSTM models (Dataset E1, configuration 1) during training and test phases

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature T_r	Configuration 1	55.06	53.58	51.93	52.13
	Configuration 2	27.92	24.67	29.44	30.54
	Configuration 3	73.42	64.14	71.76	71.45
Compressor demand power P	Configuration 1	91.22	85.35	67.05	88.25
	Configuration 2	87.08	85.54	87.10	85.77
	Configuration 3	92.58	90.23	92.01	93.08

a. Fit

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature T_r	Configuration 1	0.230	0.248	0.276	0.277
	Configuration 2	0.408	0.425	0.403	0.402
	Configuration 3	0.116	0.132	0.143	0.150
Compressor demand power P	Configuration 1	0.052	0.073	0.083	0.065
	Configuration 2	0.071	0.079	0.073	0.077
	Configuration 3	0.045	0.055	0.045	0.034

b. MAE

Table 4: Prediction performance (Fit and MAE) of 4 LSTM models (Dataset E1) in test phase for 3 configurations of inputs and outputs

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature T_r	Configuration 1	61.55	61.11	48.50	54.46
	Configuration 2	51.42	42.89	51.98	51.16
	Configuration 3	73.10	75.70	81.56	75.67
Compressor demand power P	Configuration 1	88.15	88.37	84.01	88.00
	Configuration 2	86.09	85.47	84.44	87.52
	Configuration 3	90.50	88.60	90.27	90.81

a. Fit

		Traditional LSTM	Stacked LSTM	Bidirectionnal LSTM	Convolutional LSTM
Return air temperature T_r	Configuration 1	0.269	0.280	0.371	0.335
	Configuration 2	0.385	0.436	0.381	0.388
	Configuration 3	0.203	0.172	0.139	0.162
Compressor demand power P	Configuration 1	0.075	0.081	0.099	0.077
	Configuration 2	0.091	0.085	0.076	0.070
	Configuration 3	0.068	0.071	0.053	0.066

b. MAE

Table 5: Prediction performance (Fit and MAE) of 4 LSTM models (Dataset E2) in test phase for 3 configurations of inputs and outputs