A Modular Thermal Space Coupling Approach for Indoor Temperature Forecasting Using Artificial Neural Networks

Abstract

With the increasing digitalization of buildings and the adoption of comprehensive sensing and metering networks, the concept of building digital twins is emerging as a key component in future smart and energy-efficient buildings. Such digital twins enable the use of flexible and adaptable data-driven models to provide services such as automated performance monitoring and model-based operational planning in buildings. In this context, accurate indoor temperature models are vital to ensure that the proposed operational strategies are effective, feasible, and do not compromise indoor comfort. In this work, the significance of thermal space coupling for data-driven indoor temperature forecasting is investigated by assessing and comparing the performance of an isolated and coupled Long Short-Term Memory model architecture across 70 spaces in a case study building. To construct the coupled architecture, an open-source tool is developed and presented, which allows the automated extraction of space topology from IFC-files to identify adjacent spaces. The coupled architecture is found to outperform the isolated architecture for ∼84% of the investigated spaces, with significant improvements under certain operational and climatic conditions. To account for the subset of spaces where the isolated architecture performs better, it is proposed to select between the two architectures accordingly. The demonstrated modularity and embedded adaptability of the proposed model architectures provide a sound basis for implementation in a highly dynamic building Digital Twin environment.

Key Innovations

- A novel and generic tool for automatic extraction of building topology from IFC files is presented and applied
- A thermally coupled MISO architecture for modular data-driven indoor temperature modeling and forecasting is presented and assessed

Practical Implications

The presented tool and architectures provides an automated, scalable, and accurate method for modeling and forecasting indoor temperature, enabling decision makers to evaluate how different operational strategies affect the indoor comfort levels of a building.

Introduction

Energy consumption in buildings is necessary to sustain proper air quality and adequate thermal comfort for occupants. However, the building sector is currently one of the largest energy-consuming sectors with major potential for improvements in energy efficiency and flexibility. With this as the driving force, buildings are currently seeing a fast-paced digitalization in the pursuit of Internet of Things (IoT)-integrated smart buildings (Jia et al. (2019)). In this context, data-driven modeling methods have received considerable attention for numerous applications, due to the promising potential for automation and scalability. The adaptability of data-driven approaches has been demonstrated through various applications, e.g. in forecasting of indoor air quality (Wei et al. (2019)), indoor temperature (Alawadi et al. (2020)), building occupancy (Jin et al. (2021)), energy loads (Amasyali and El-Gohary (2018)) as well as for fault detection and diagnosis (FDD) (Mirnaghi and Haghighat (2020)).

Indoor temperature is considered one of the key indicators of indoor comfort and accurate prediction models are therefore crucial for exploring different operational strategies while maintaining proper indoor comfort. Recently, black-box modeling and especially Artificial Neural networks have gained increased attention for indoor temperature modeling with recent studies suggesting a potential performance gain in considering the thermal coupling between spaces in buildings using large monolithic Multiple Input Multiple Ouput (MIMO) architectures (Mtibaa et al. (2020); Fang et al. (2021)). However, the performance of such architectures varies greatly with the specific topology and operational patterns of the modeled building and can in some cases be worse than the simpler isolated Multiple Input Single Ouput (MISO) model architectures (Mtibaa et al. (2020)). Furthermore, relying on one large black-box model for modeling the whole building makes it difficult to understand and rectify the model, in case of poor model performance. Therefore, we propose a different MISO architecture, where each space is modeled individually and only locally relevant thermal interaction between spaces is considered.

The benefits of this are twofold: 1) Each space is modeled separately, adding much greater modularity

and flexibility, compared to the mentioned monolithic MIMO architecture. For instance, using the proposed MISO architectures, some spaces can be modeled as isolated while other spaces with non-negligible heat transfer to adjacent spaces can be modeled using a coupled architecture. 2) More training data is available due to the lower amount of inputs per model, which can potentially improve performance (this is further discussed in section).

For the proposed coupled architecture to be feasible for practical implementation, an automated tool, Ifc2Graph (Bjørnskov (2022)), has been developed. Based on the open Industry Foundation Classes (IFC) standard, the tool analyses the building geometry to identify adjacent spaces. The two architectures are tested on 70 spaces, highlighting the performance of the two architectures in a realistic building setting. The considered spaces are occupied and include radiators for space heating, demand-controlled ventilation, and automated shading.

The presented work is in line with the ongoing digitalization of the building sector and the need for smarter buildings. The modularity and flexibility of the proposed model architectures ensure seamless integration with future building Digital Twins, where accurate and flexible indoor temperature models play a crucial role in safely exploring and recommending optimal building operation strategies without compromising indoor comfort. The presented work is carried out as part of the international research project 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle', aiming to design, develop, and demonstrate a holistic Digital Twin platform allowing coordinated decision support over the whole building life cycle. The proposed Digital Twin will provide the three following services: i) Flexible and effective data collection and management employing open standard context information models; ii) Smart facility management through performance monitoring and automated commissioning; and iii) Planning support through model-based informed decision-making.

Case study

In this work, a Danish university building is considered as a case study for testing the two model architectures. The building was commissioned in 2015 and was one of the first buildings in Denmark to comply with the Danish building class 2020. It has four stories, including a basement with technical rooms, storage, and installations. The ground floor, first floor, and second floor are mainly devoted to teaching and are composed of classrooms, offices, corridors, and study zones. An image of the building is shown in Figure 1.

The indoor climate is regulated through four balanced mechanical ventilation systems, each with a capacity of 35,000 m3/h, while the heating source of the building is district heating, supplying air-loop heating coils as well as radiators distributed across all four stories. Most spaces are well-instrumented with temperature, CO_2, humidity, PIR, and lux sensors. Furthermore, all radiators are equipped with thermostatic valves that control the indoor temperature by actuating the valve position. To ensure a proper indoor climate, the BMS employs Demand Controlled Ventilation (DCV) to regulate the airflows supplied to the different zones in the building. Specifically, for each space, the damper positions are actuated to prevent the measured CO_2-concentration from exceeding 600 ppm. To avoid overheating in the summer months, all outside-facing windows are equipped with shades that are controlled by the BMS. With the large scheme of sensors and meters instrumentation, the building is used as a living lab to carry out investigations and testing of various innovative applications and research-based solutions.

Figure 1: The considered case study university building (Jradi et al. (2017)).

Local weather data on hourly resolution in the form of outdoor temperature, longwave solar irradiation, and shortwave solar irradiation is available for the site in hourly resolution. In Table 1, an overview of the relevant data is given including a shorthand notation, which will be used in the following sections. For the 8 types of data, we distinguish between data collected at space-level and data collected at building level.

Table 1: Overview of relevant data collected from the building.

Notation	Type	Space level (10 min)	Building level (1 hour)
T_z	Indoor temperature	X	
C_z	CO_2-concentration	X	
u_v	Radiator valve position	X	
u_d	Air diffusers damper position	X	
u_{sh}	Shades position	X	
T_o	Outdoor temperature		X
Φ_L	Longwave solar irradiation		X
Φ_S	Shortwave solar irradiation		X

Methodology

In this section, the methodology used in this work is presented. Overall, two model architectures are investigated for the prediction of indoor temperature. First, the model architectures and chosen inputs for each of the model types are presented. Following, it

is explained how the overall space topology can be automatically extracted from a Building Information Model (BIM). It is then described how the collected data has been preprocessed and how the resulting two types of data sets differ. Finally, the training and testing methodology is presented.

Model architecture

In this work, a specific sub-category of ANNs called Recurrent Neural Networks (RNN) is used. The fundamental property of RNNs, that distinguishes them from other ANN variants such as the Multi-Layer Perceptron (MLP), is the inclusion of a system state that changes over time. This additional notion of time makes RNNs especially useful for the modeling of dynamic systems. In fact, the mathematical form of vanilla RNNs is essentially a discretized form of Delay Differential Equations (DDE) (Sherstinsky (2020)). The Long Short-Term Memory (LSTM) model, developed by Hochreiter and Schmidhuber (1997), is a specific type of RNN, mainly designed to solve the *Vanishing gradient problem* that was encountered with the vanilla RNN model. The main contribution of the LSTM model is the addition of gated units that control how information is added or removed from the system state. For the LSTM model, the state is described by two state vectors, the cell state $c \in \mathbb{R}^n$ and the hidden state $h \in [-1, 1]^n$, where n is a hyperparameter that describes the size of the state vectors, i.e. how much information that can be transferred between time steps.

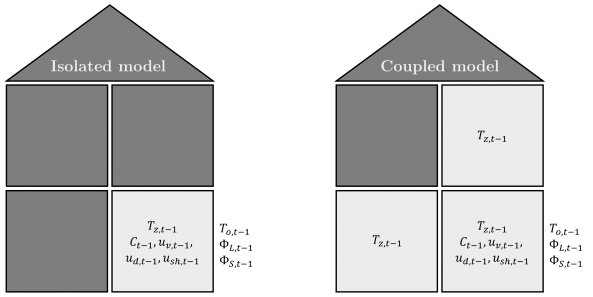

Figure 2: Inputs for the isolated and coupled architectures. The coupled architecture includes indoor temperatures from adjacent spaces as input.

As already mentioned, two types of models are investigated in this work. These models will in the remaining sections be referred to as the *isolated model* and the *coupled model*. The isolated model is based on previous work of the authors (Bjørnskov et al. (2022)), where the model inputs were determined based on well-known heat transfer mechanisms, including radiation, conduction through external surfaces, internal heat gains, ventilation, and heat added by radiators. The inputs of the coupled model consider the same heat transfer mechanisms. However, in addition to that, it also considers heat transfer through the non-external bounding surfaces, by including the indoor temperature of the adjacent spaces as input. An overview of the two model types is shown in Figure 2. As seen, the only difference between the two model types is the inclusion of adjacent space temperatures as input in the coupled model. The network architecture of the two investigated models is seen in Figure 3. Both the isolated model and the coupled model consist of two sequential LSTM networks, A and B. Network A receives the inputs shown in Figure 2, i.e. it receives one set of inputs for the isolated model and another set of inputs for the coupled model. In addition, Network A receives as input the state vectors from the previous time step $c_{A,t-1}$ and $h_{A,t-1}$ and outputs the updated state vectors $c_{A,t}$ and $h_{A,t}$. The size n of the state vectors c_A and h_A will be referred to as n_A.

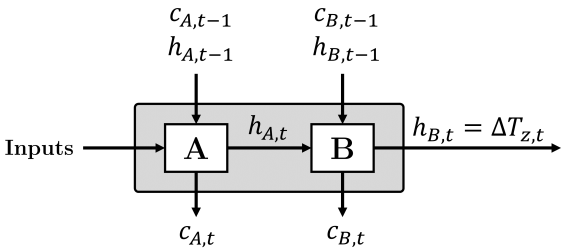

Figure 3: Model architecture for both the isolated and coupled model consisting of the sequential LSTM networks A and B. The model output is the predicted temperature change of the modeled space during a given time step.

Network B receives as input the updated hidden state $h_{A,t}$ from network A along with its own state vectors $c_{B,t-1}$ and $h_{B,t-1}$ from the previous time step. Network B outputs the updated state vectors $c_{B,t}$ and $h_{B,t}$, both with size $n = 1$. $h_{B,t}$ is treated as the model output which is fitted to the chosen prediction target during training. As seen in Figure 3, the chosen prediction target of the model is the temperature change of the modeled space during a given time step $\Delta T_{z,t}$. Through earlier investigations by the authors, it was found that this configuration (rather than directly predicting temperature) generalized better for custom control scenarios, e.g. when the model is implemented in a setpoint control loop (Bjørnskov et al. (2022)). The use of this configuration for indoor temperature predictions is later explained in section .

Automated extraction of space topology

For a coupled model to be feasible in larger buildings, an efficient and automated method of identifying adjacent spaces is required. With the significant adoption rate of Building Information Modeling (BIM), automatic extraction of geometrical building data for the use in Building Energy Modeling (BEM) tools is an active research area (BIM2BEM) and various algorithms have been developed for this purpose (Treeck and Rank (2006); Rose and Bazjanac (2013); Jones et al. (2013); Lilis et al. (2016)). Traditional

BEM tools require the paring of *space boundaries*, each associated with thermal heat transfer between two spaces. However, such detailed surface paring is not necessary for the application presented in this work, where the topology of adjacent spaces is sufficient. Furthermore, no accessible open-source tools have been identified in the literature that can fulfill this purpose. Therefore, the authors have chosen to develop such an algorithm based on the Industry Foundation Classes (IFC) standard (ISO (2018)), one of the leading BIM standards (Edirisinghe and London (2015)).

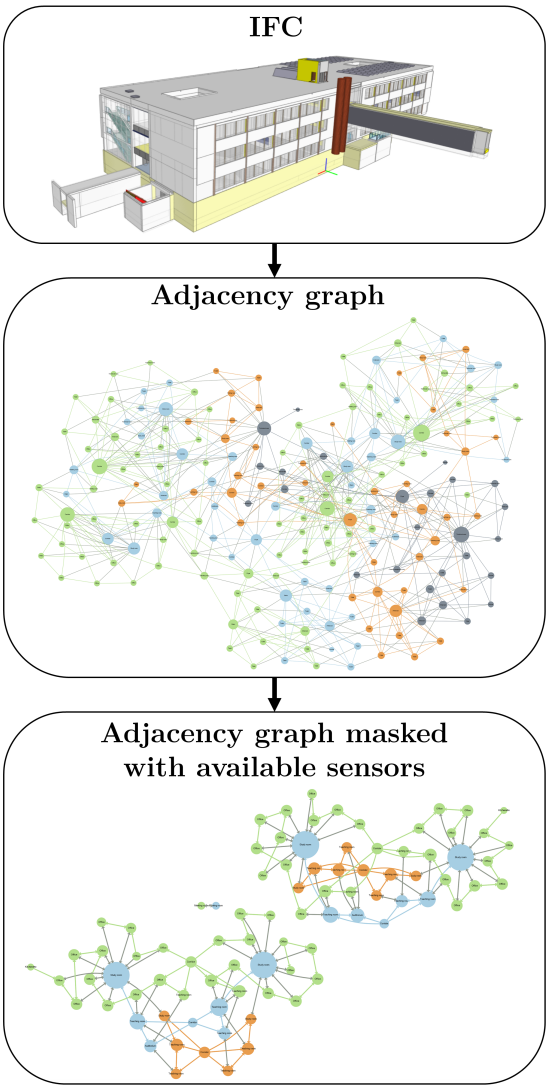

Figure 4: IFC file of the case study building (top), the derived graph that represents the adjacency of spaces (middle), and the adjacency graph masked with available sensors (bottom). The node colors represent the story levels with grey, orange, blue, and green representing the basement, ground, first floor, and second floor, respectively. The node size is scaled with respect to the number of connected edges, i.e. the number of adjacent spaces.

Generally, the IFC format offers detailed information on building geometry and design data. Typically, all geometrical information relevant for modeling buildings is grouped under `IfcSpace` entities. The geometry of the building can thus be extracted on a space-by-space level and used to determine which spaces share boundaries. For this purpose, a python package called Ifc2Graph has been developed and made available as an open-source tool on GitHub (Bjørnskov (2022)). The developed tool has been tested on multiple IFC files to ensure that it generalizes beyond the case study considered in this work.

An image of the IFC file containing the architecture of the case study building is seen in Figure 4 (top). Using this IFC as input for the algorithm, yields the graph shown in Figure 4 (middle). Here, nodes represent spaces, one for each `IfcSpace` entity in the IFC file, while edges represent adjacency of the two connected spaces. The node colors show the story levels with grey, orange, blue, and green representing the basement, ground, first, and second floor, respectively. The size of each node has been scaled in correspondence to the number of connected edges, i.e. the number of adjacent spaces.

In this work, we concentrate on spaces with the energy balance described in section . Therefore, only a subset of the shown nodes with all five space-level data types described in Table 1 is selected for further modeling. This yields the masked adjacency graph in Figure 4 (bottom). The masked graph is obtained by removing all nodes from the adjacency graph that does not have all five space-level sensor types described in Table 1. Spaces with temperature measurements that are adjacent to one or more nodes fulfilling the above criteria are also kept in the graph. Hence, the nodes in this graph represent spaces that are either modeled or are supplying indoor temperature data to modeled spaces.

The case study building is almost symmetrical, meaning that comparing the north-south and west-east sections of the building yields an almost identical number of spaces and space types in each section. As seen, this is reflected in the masked adjacency graph, which consists of two large clusters that are almost identical in structure. The clusters are not connected because the central hallway of the building, dividing the building into its north and south sections has no installed temperature sensors. A small third cluster is also shown, consisting of two spaces. Again, none of these two rooms are adjacent to any space with installed temperature sensors.

Data preprocessing

To compare the performance of the isolated space models and the coupled space models, two types of data sets have been constructed, each with the inputs described in section . For each space, the inputs have been sampled on a 10 minutes interval in 24-hour sequences with 144 time steps, based on actual

sensor measurements from the building during a two-year period from January 1st 2018 to December 31st 2019.

For the coupled architecture, there can be certain drawbacks to introducing additional inputs to a given model. For instance, consider Figure 5 where time steps for three inputs and one output are shown during a certain period after preprocessing of the data. Green represents the availability of data and red represents the absence of data for each respective time step. The absence of data could be due to many reasons, e.g. missing values, removed outliers, frozen sensor readings etc. Note that for the actual data set, linear interpolation is performed for gaps of 12 time steps or less. The gaps of 1 to 3 time steps in the example would therefore not be present after preprocessing.

To find which time steps can be used for constructing training sequences, the intersection of data available across all inputs and outputs is found. Using this intersection, unique sequences with the desired step length can then be identified. For instance, if the sequence length is 3, then a data set of 8 sequences can be constructed for the given example. Therefore, introducing additional model inputs can potentially lower the number of training sequences, e.g. by adding adjacent space temperatures as model inputs.

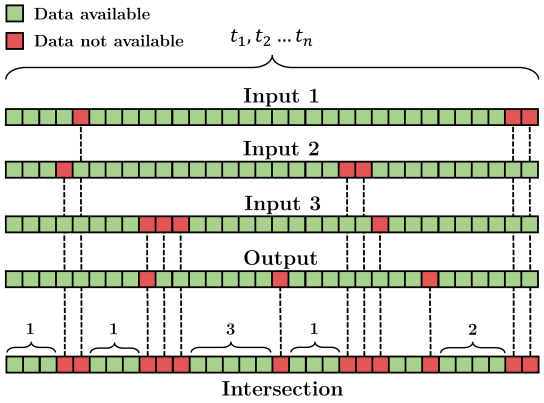

Figure 5: Example of intersection across three inputs and one output to find suitable time steps in the collected data for the construction of training, validation, and testing sequences.

The distribution of data sequences for both data sets after preprocessing across the different seasons is shown in Figure 6. Here, the sequence count across all spaces has been sorted for each season and data set, as represented by the colored area. Furthermore, these distributions have been summarized by the shown box plots. The sequence count varies notably between seasons. As seen, the inclusion of adjacent spaces in the model inputs lowers the sequence count across all seasons with the winter season having the most significant data loss. For the coupled datasets, the sequence count distribution is effectively shifted down for all seasons.

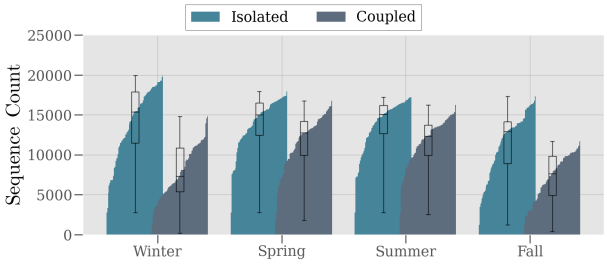

Figure 6: Sequence count per space distributed over the seasons for the two preprocessed data sets.

Training and testing method

For each space, the two data sets described in section , have been split into training, validation, and testing data sets with the splits [1/3, 1/6, 1/6], respectively. To ensure that seasonality and varying building operation is reflected properly in each data set, data from each month has been distributed in proportion to the mentioned data splits in the training, validation, and testing data sets. In this process, it has been ensured that sequences between training and test data sets do not overlap to avoid data leakage.

All models have been trained using the python library Pytorch and Stochastic Gradient Descent (SDG) as optimizer with the Mean Square Error (MSE) loss function. This work considers a one-fit-all strategy where the same set of hyperparameters is used across all models. Although hyperparameter tuning performed on a model-by-model basis would likely improve performance, it would also increase computational demands and training time considerably. Through initial testing, a batch size of 32, a learning rate of 0.1, a momentum of 0.9, and a state size $n_A = 20$ was found to give satisfactory performance across all spaces. All models have been trained using an early-stopping technique to avoid overfitting. This is achieved by monitoring the loss on the validation data set while saving copies of the model instances at specific gradient step intervals during training. When the performance stops to improve or starts to degrade, the model instance with the lowest recorded validation loss is kept while the rest is discarded.

In earlier studies, the performance of indoor temperature models is often evaluated only for one-step-ahead predictions (Marvuglia et al. (2014)). However, if the step size is small, e.g. between 1-10 minutes, the obtained performance metrics are not sufficient to assess the model performance for longer forecasting horizons such as hours or days, due to the large thermal inertia and slow-moving dynamics of buildings. Therefore, to reflect model performance across longer horizons, this work considers a forecasting horizon of 24 hours. As explained in section , the models predict temperature change instead of the actual space temperature

on a 10-minute step interval. Therefore, to measure the model performance in terms of temperature prediction, the model is implemented in a closed-loop configuration as shown in Figure 7.

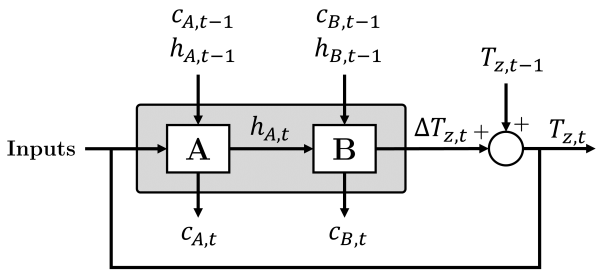

Figure 7: Closed-loop configuration for indoor temperature forecasting.

In Figure 7, the current indoor temperature change $\Delta T_{z,t}$ predicted by the model is added with the previous temperature value $T_{z,t-1}$ to obtain the current indoor temperature $T_{z,t}$. This newly obtained temperature is then fed back as input to the model, thus creating a loop. This process can be repeated for an arbitrary number of steps, as long as the defined inputs are available.

Results and Discussion

To assess and compare the performance of the isolated and coupled architectures, the closed loop configuration described in section is employed on the constructed test data sets that contain 24-hour sequences of 144 steps for all four seasons. For each of these sequences, predictions are thus obtained for the indoor temperature, which can be compared to the actual measured temperature. For this purpose, the Mean Absolute Error (MAE) is used as a metric to evaluate the performance of the trained models. Furthermore, to see how the prediction accuracy varies with different operational and ambient conditions, the MAE has been calculated on a seasonal basis for all models. The results are shown in Figure 8 where the colored areas show the sorted seasonal MAE across all spaces for both model architectures and boxplots have been added to summarize these distributions.

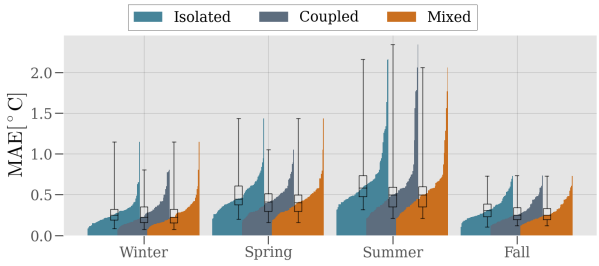

Figure 8: Mean Absolute Error per space model on the testing data set distributed over the seasons for the two model architectures.

As seen in Figure 8, the coupled model architecture yields better predictions for spring, summer, and fall with a lower median, first quartile, and third quartile. Especially for the summer season, a significant difference in performance can be observed between the two architectures. The building has no installed cooling equipment, and the building is thus mostly unconditioned during summer. For an unconditioned building, the energy balance of each space is typically dominated by heat transfer through the surrounding walls and windows, i.e. by temperature differences. Furthermore, since the indoor temperatures for this scenario are no longer forced to follow the specified setpoints, the temperature variance across the building is also greater. This explains the improved performance of the coupled model, where boundary temperatures are provided as inputs.

For the winter season, where the building is conditioned and the indoor temperature is mostly uniform, the isolated architecture performs similar to the coupled architecture with a higher first quartile, but a lower third quartile. This indicates that 25% of the best-performing isolated models perform worse than the equivalent fraction for the coupled models. However, it also follows that 25% of the worst-performing isolated models perform better than the equivalent fraction for the coupled models.

Contrary to the suggested monolithic MIMO architectures in recent studies (Mtibaa et al. (2020); Fang et al. (2021)), the presented MISO architectures allow for a modular modeling approach where the model architecture and performance can be calibrated on a space-by-space level. Spaces with strong thermal coupling to adjacent spaces can thus be modeled with the coupled architecture, while spaces having negligible thermal interaction with their adjacent spaces can be modeled using the isolated architecture.

This is demonstrated in Figure 8, where a *Mixed* category is shown. Here, either the isolated or coupled model has been selected, based on their performance on the validation data set, resulting in the selection of 11 isolated models and 59 coupled models. This slightly improves most aspects of the performance distribution across all seasons.

To validate the trained models on a qualitative basis, three of the coupled space models, representing a classroom, a study zone, and an office, have been chosen for further investigations. For this purpose, a 24-hour winter sequence from the test dataset is used as input with the closed-loop configuration described in section . The results are shown in Figure 9. Here, the individual inputs for each of the three models are shown in the six plots to the left while the common weather input is shown on the rightmost plot. As seen from the weather data, outdoor temperatures range between -5 to 5 °C, while irradiation levels are generally low. The three bottom plots on the left show

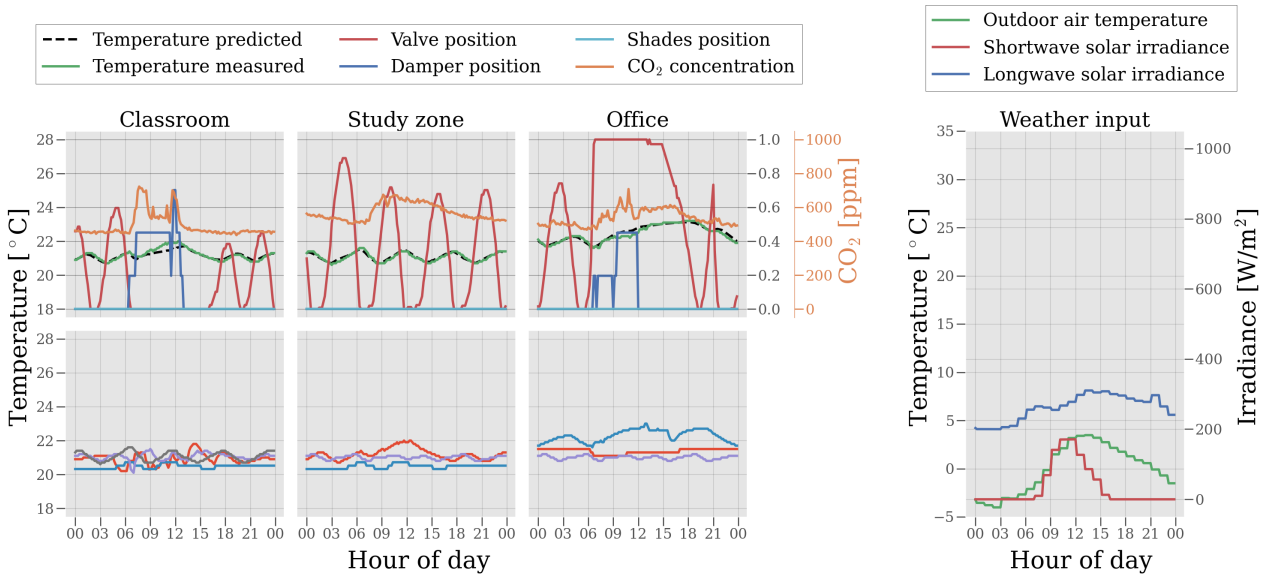

Figure 9: 24-hour forecasting for three spaces during a winter period using the coupled model architecture.

the input temperatures of the adjacent spaces, which are mostly within the range of 20 to 23 °C. The classroom includes four such inputs while the study zone and the office include three. The three top plots on the left show the remaining model inputs as well as the measured and predicted indoor temperature for each space. As seen, the model predictions agree very well with the measured temperatures throughout all 24 hours with only minor errors. During the winter period, the radiator valve position appears to be the most influential input with respect to the indoor temperature. To reach the specified setpoint temperature for each space, the valve position is opened when the measured temperature falls below the setpoint and is closed when the temperature surpasses the setpoint. For the classroom and the study zone, this results in a periodically fluctuating valve position and indoor temperature, which is well-captured by the coupled model architectures. Although the valve position of the space heater has the largest influence on indoor temperature, the impact of the dampers, regulating the airflows, can also be observed.

For the Classroom and the Office, the dampers are opened between hours 06 and 12 as a response to the increasing CO_2-concentration, which in periods exceeds 600 ppm. The isolated impact of increased air exchange on the indoor temperature is visible for the classroom, where the radiator valve is closed during this period. Here, the model predicts a slightly lower temperature increase compared with the actual measured temperature. At approximately 12, where the dampers are closed, the temperature starts to decrease until at approximately 16, when the valve position is again opened.

Conclusion

Previous studies have highlighted that the benefit of coupled MIMO models largely depends on the type of building and its operation and that the effectiveness of isolated MISO and coupled MIMO architectures must be evaluated on a building-by-building basis. In this work, it was investigated how thermal coupling between adjacent spaces affects the prediction accuracy of indoor temperature by comparing an isolated and a coupled LSTM MISO architecture across 70 spaces in a case study building. To construct the coupled models, a novel open-source tool Ifc2Graph was presented and used to automatically extract the space topology from an IFC-file of the case study building. For each space and architecture, data from various sensor and meter types across a two-year period was used to train and test the developed models. It was found that the coupled architecture improved overall performance for majority of spaces (∼84 %), with significant improvements during the summer season, where the case study building was mostly unconditioned.

The modularity of the presented architectures allows the modeling of spaces with strong thermal coupling to adjacent spaces using the coupled architecture, while spaces with insignificant heat transfer to adjacent spaces can be modeled using the isolated architecture. In the broader context of future smart buildings and building digital twins, this added modularity will allow for increased adaptability to different building types, climate conditions, and data availability, while ensuring that accurate indoor temperature predictions are obtained. With indoor temperature being a key-indicator for the comfort-levels of occupants in buildings, the presented approach can be used to obtain accurate indoor temperature models,

that can help asses the feasibility of different operational strategies in buildings. For such an application to be reliable, future studies should investigate how the uncertainty of each prediction can be quantified and how well the models generalize to unseen operational conditions. Furthermore, the proposed method should be tested in environments with varying quantities of data to asses how this affects the performance, e.g. to simulate implementation in a newly built building.

Acknowledgment

This work is carried out under the 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle' project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

References

Alawadi, S., D. Mera, Fernández-Delgado, F. Alkhabbas, Carl, M. Olsson, P. Davidsson, and S. Alawadi (2020, 01). A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings. *Energy Systems*.

Amasyali, K. and N. M. El-Gohary (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews 81*, 1192–1205.

Bjørnskov, J. (2022). Ifc2graph. `https://github.com/JBjoernskov/Ifc2Graph`, Last Accessed: 06-17-2022.

Bjørnskov, J., M. Jradi, and C. Veje (2022, 07). A fully automated and scalable approach for indoor temperature forecasting in buildings using artificial neural networks. In *Building Simulation Applications BSA 2022*.

Edirisinghe, R. and K. London (2015, 10). Comparative analysis of international and national level bim standardization efforts and bim adoption.

Fang, Z., N. Crimier, L. Scanu, A. Midelet, A. Alyafi, and B. Delinchant (2021). Multi-zone indoor temperature prediction with lstm-based sequence to sequence model. *Energy and Buildings 245*, 111053.

Hochreiter, S. and J. Schmidhuber (1997). Long short-term memory. *Neural Computation 9*, 1735–1780.

ISO (2018). International organization for standardization, "iso 16739-1:2018". `https://www.iso.org/standard/70303.html.`, Last Accessed: 06-07-2022.

Jia, M., A. Komeily, Y. Wang, and R. S. Srinivasan (2019). Adopting internet of things for the development of smart buildings: A review of enabling technologies and applications. *Automation in Construction 101*, 111–126.

Jin, Y., D. Yan, A. Chong, B. Dong, and J. An (2021). Building occupancy forecasting: A systematical and critical review. *Energy and Buildings 251*, 111345.

Jones, N., C. McCrone, B. Walter, K. Pratt, and D. Greenberg (2013, 08). Automated translation and thermal zoning of digital building models for energy analysis. pp. 202–209.

Jradi, M., F. C. Sangogboye, C. G. Mattera, M. B. Kjærgaard, C. Veje, and B. N. Jørgensen (2017). A world class energy efficient university building by danish 2020 standards. *Energy Procedia 132*, 21–26. 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway.

Lilis, G., G. Giannakis, and D. Rovas (2016, 12). Automatic generation of second-level space boundary topology from ifc geometry inputs. *Automation in Construction 76*.

Marvuglia, A., A. Messineo, and G. Nicolosi (2014). Coupling a neural network temperature predictor and a fuzzy logic controller to perform thermal comfort regulation in an office building. *Building and Environment 72*, 287–299.

Mirnaghi, M. S. and F. Haghighat (2020). Fault detection and diagnosis of large-scale hvac systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings 229*, 110492.

Mtibaa, F., K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne, and M. Cheriet (2020, 12). Lstm-based indoor air temperature prediction framework for hvac systems in smart buildings. *Neural Computing and Applications 32*.

Rose, C. and V. Bazjanac (2013, 04). An algorithm to generate space boundaries for building energy simulation. *Engineering with Computers 31*.

Sherstinsky, A. (2020, 03). Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena 404*, 132306.

Treeck, C. and E. Rank (2006, 01). Dimensional reduction of 3d building models using graph theory and its application in building energy simulation. *Engineering with Computers An InternationalJournal for Simulation-Based Engineering 23*, 109–122.

Wei, W., O. Ramalho, L. Malingre, S. Sivanantham, J. Little, and C. Mandin (2019, 06). Machine learning and statistical models for predicting indoor air quality. *Indoor Air 29*.