Jornadas de Automática

# Transfer Learning for Time-series Forecasting of Greenhouse Microclimate

Loer, Paul[a,*], Daniels, Annalena[a], Fink, Michael[a], García-Mañas, Francisco[b], Wollherr, Dirk[a], Rodríguez, Francisco[c]

[a] *Chair for Automatic Control at the Technical University of Munich, Germany.*
[b] *University of Seville, Department of Systems Engineering and Automation, 41092 Seville, Spain.*
[c] *University of Almería, Department of Informatics, CIESOL, ceiA3, 04120 Almería, Spain.*

**To cite this article:** Loer, Paul, Daniels, Annalena, Fink, Michael, García-Mañas, Francisco, Wollherr, Dirk, Rodríguez, Francisco. 2025. Transfer Learning for Time-series Forecasting of Greenhouse Microclimate. Jornadas de Automática, 46. https://doi.org/10.17979/ja-cea.2025.46.12186

**Abstract**

This paper presents the adaption of a data-driven model of the microclimate of a first greenhouse (of 877 $m^2$) to a second, different greenhouse (of 1900 $m^2$) by using transfer learning. The temporal evolution of air temperature and relative humidity inside the first greenhouse were modeled using a transformer-based model. The transformer was trained and validated with an 81-day dataset from the first greenhouse, for multi-step prediction of the microclimate. Subsequently, the same transformer-based model was used with a different 48-day dataset from the second greenhouse. Transfer learning was then used to fine-tune the weights and biases of the transformer. Results show that, by using transfer learning, only two days of data were necessary to train the transferred model. This also shows that the initial transfomer generalized well and learned basic greenhouse climate dynamics that only need to be adapted slightly when the greenhouses specifications change. Therefore, transfer learning is presented as a method that can facilitate the reuse or adaptation of neural-based predictive controllers between different greenhouses.

*Keywords:* Transfer Learning, Greenhouse Modeling, Transformers, Time-series Prediction, Multi-step Prediction

## 1. Introduction

Greenhouses are essential for sustainable, high-efficiency food production because they allow crops to be cultivated under controlled environmental conditions. Maintaining a stable microclimate, specifically temperature and humidity, is critical for achieving consistent yields and efficient resource use, as these factors directly affect plant health and growth (Rodríguez et al., 2015).

Accurate modeling of the greenhouse microclimate is a key aspect for implementing predictive control systems that dynamically adjust heating, ventilation, and shading mechanisms in response to changing internal and external conditions. However, greenhouse climate dynamics are highly variable and influenced by structural design, orientation, vent placement, crop type, and seasonal changes. External weather factors, such as radiation, temperature, humidity, and wind, further complicate the system by introducing significant nonlinearities, making accurate modeling a difficult task. Although physics-based models use thermodynamic equations to provide a principled foundation, they are often difficult to calibrate and do not generalize well across different greenhouse configurations. Furthermore, collecting the necessary data to calibrate these models for each new setup is time-consuming and inefficient. Data-driven methods offer a promising alternative by learning dynamic patterns directly from observations (Patil et al., 2023). For instance, in (Guo and Feng, 2024) and (Lin et al., 2024), various types of Long Short-Term Memories (LSTMs) were successfully used to learn and predict the temporal behavior of greenhouse systems. Also, Fink et al. (2025) use a feedforward Neural Network (NN) to model the greenhouse microclimate dynamics for a Learning-based Model Predictive Control approach.

However, training such data-driven models typically requires large datasets, which are not always available for every application. In this context, transfer learning is a practical solution because it allows pre-trained models to be adapted to new greenhouses by fine-tuning with a minimal amount of additional data to achieve the same prediction errors as with a

[*] Corresponding author: paul.loer@tum.de

single large dataset (Ismail Fawaz et al., 2018). This enables faster and more scalable deployment of predictive systems as collecting datasets on greenhouses is bound to time and take several weeks. For example, a model trained with data measured at one specific greenhouse could be adapted to simulate the microclimate of a different greenhouse by a transfer learning approach. In recent years, some studies have been published evaluating the potential of applying transfer learning in greenhouses, both for modeling and control approaches (Moon and Eek Son, 2021; Zhao et al., 2022). Additionally, new model architectures based on NNs, such as transformer-based models, have been applied to greenhouses with promising results for modeling the microclimate (Lee et al., 2025).

This work introduces a framework for modeling the greenhouse microclimate using transformer-based neural networks. Results show that this approach can accurately predict internal air temperature and humidity based on external weather conditions and the vent opening percentage, which are key variables for climate control in Mediterranean greenhouses when using natural ventilation. By fine-tuning pre-trained models on small datasets from new greenhouses, transfer learning significantly reduces the data and time required to develop predictive models suitable for control applications. It effectively accounts for variability in greenhouse structures, crop types, and seasonal conditions, thereby improving generalization. This makes it feasible to apply machine learning models in real-world scenarios without the need for prolonged data collection and tuning periods.

The remainder of the paper is structured as follows. Section 2 describes the materials and methods, including datasets measured in two Mediterranean greenhouses (located in different geographical areas with some climatic differences), transformer-based models, and the transfer learning method. Section 3 presents the results and discussion. The work ends with a summary of conclusions and future work in Section 4.

## 2. Materials and Methods

### *2.1. Description of Greenhouse Datasets*

The transformer-based model later presented in Section 2.4 needs to be trained by greenhouse data, considering that the model predictions are the states $\hat{\boldsymbol{x}}_{i|k}$ of the greenhouse (i.e., the climate variables of interest). This work utilizes two different datasets of two greenhouses as the source and target domain and task, as explained in the following sections.

#### *2.1.1. Source Dataset*

The source dataset contains data acquired by several sensors from an Almería-type greenhouse (877 $m^2$) located at "Las Palmerillas" Experimental Station of the Cajamar Foundation in El Ejido, Almería, Spain (see Figure 1a). For natural ventilation, it is equipped with sevens vents in total, five being located on the roof and two on the north and south sidewalls, respectively. The data from this greenhouse were recorded from October 10, 2020, to December 29, 2020, by taking measurements of all climate variables and vent positions (from 0% to 100% of vent opening) every 30 s, which results in 233 381 data points recorded over a period of 81 days. The dataset is presented in Figure 2.

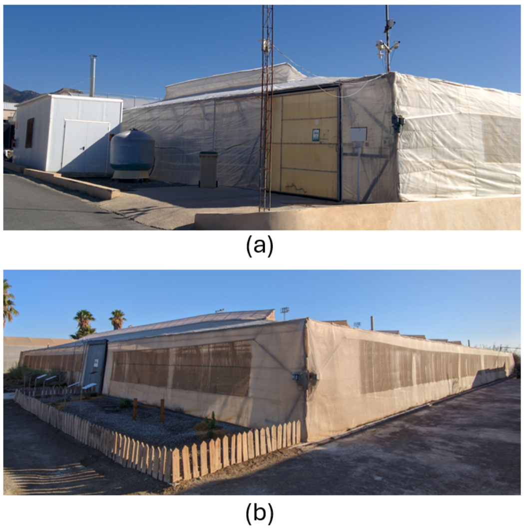

Figure 1: Experimental greenhouses. The source in (a) and the target greenhouse in (b).

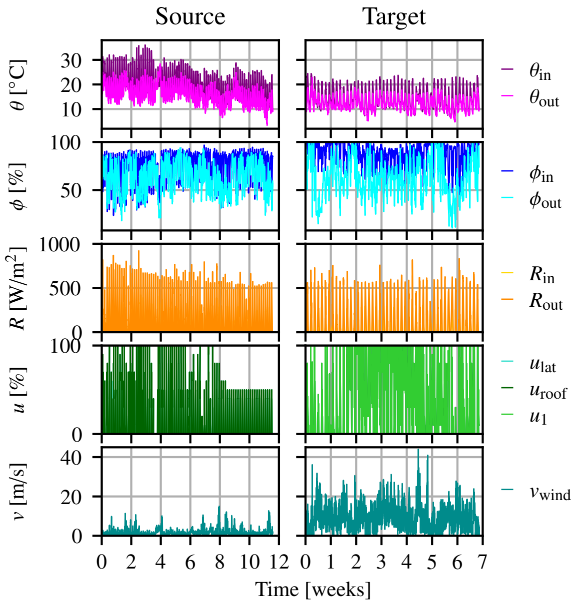


Figure 2: Plot of source and target dataset. See Section 2.2 for the variables.

#### *2.1.2. Target Dataset*

The second dataset, the target dataset, stems from another Almería-type greenhouse (1900 $m^2$) located at the AgroConnect.es research facilities and IFAPA center in La Cañada, Almería, next to the University of Almería, Spain (see Figure 1b). It was also recorded in winter, from December 6, 2024, to January 22, 2025. Its data are presented in Figure 2. While the measured climate variables were the same, this greenhouse has more vents for natural ventilation: six roof vents and seven side vents. All vents can be controlled separately, but they were synchronously controlled in the recorded dataset.

### *2.2. Greenhouse Modeling Framework*

Figure 3 shows a schematic of the greenhouses as a reference framework for the modeling procedure later explained in Sections 2.3 and 2.4. Therefore, the greenhouse state space is defined as follows. At time step $k$, the state is described by $\mathbf{x}_k = [\theta_{\mathrm{in}}, \phi_{\mathrm{in}}]$ with the inside temperature $\theta_{\mathrm{in}}$ and humidity $\phi_{\mathrm{in}}$. The lateral and roof vents form the control input with $\mathbf{u}_k = [u_{\mathrm{roof}}, u_{\mathrm{lat}}]$. The disturbances of the system are in this case the outside weather variables of the greenhouse, which are defined as $\mathbf{p}_k = [\theta_{\mathrm{out}}, \phi_{\mathrm{out}}, R_{\mathrm{in}}, R_{\mathrm{out}}, v_{\mathrm{wind}}]$ with the outside temperature $\theta_{\mathrm{out}}$ and humidity $\phi_{\mathrm{out}}$, the in- and outside radiation $R_{\mathrm{in}}$ and $R_{\mathrm{out}}$, and the outside wind speed $v_{\mathrm{wind}}$. During the measurements, a tomato crop was growing in the two greenhouses described in Section 2.1.

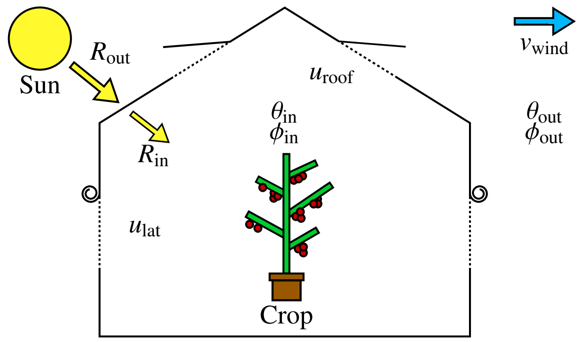


Figure 3: Schematic of the greenhouse modeling framework.

### 2.3. Datasets Preparation for Model Training

When training a NN, samples from the dataset are provided as input. Each sample at step $k$ is defined as the vector $\boldsymbol{s}_k = [\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{p}_k]^{\mathrm{T}}$. The number of prediction steps typically depends on the intended application; in this case, a multi-step prediction is employed, as forecasts over a time span of several hours are required. Accordingly, a sequence of past samples $\boldsymbol{S}_k^-$ and known future samples $\boldsymbol{S}_k^+$ are fed to the NN for prediction, where $\boldsymbol{S}_k^-$ and $\boldsymbol{S}_{k+1}^+$ are defined as

$$\boldsymbol{S}_k^- = \left[\boldsymbol{s}_{k-m\cdot(l-1)}^{\mathrm{T}}, \boldsymbol{s}_{k-m\cdot(l-2)}^{\mathrm{T}}, \dots, \boldsymbol{s}_k^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{1}$$

$$\boldsymbol{S}_{k+1}^+ = \left[\widetilde{\boldsymbol{s}}_{k+1}^{\mathrm{T}}, \dots, \widetilde{\boldsymbol{s}}_{k+1+m\cdot(l-2)}^{\mathrm{T}}, \widetilde{\boldsymbol{s}}_{k+1+m\cdot(l-1)}^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{2}$$

with the future known samples defined as $\widetilde{\boldsymbol{s}}_k = [\widetilde{\boldsymbol{u}}_k, \widetilde{\boldsymbol{p}}_k]^{\mathrm{T}}$. These sequences can then be used to predict the future states $\boldsymbol{X}_{k+1}^+$ of the greenhouse microclimate with the prediction model $\boldsymbol{G}_{\mathrm{NN}}$ as in $\hat{\boldsymbol{X}}_{k+1}^+ = \boldsymbol{G}_{\mathrm{NN}}\left(\boldsymbol{S}_k^-, \boldsymbol{S}_{k+1}^+\right)$ with

$$\hat{\boldsymbol{X}}_{k+1}^+ = \left[\hat{\boldsymbol{x}}_{k+1}^{\mathrm{T}}, \dots, \hat{\boldsymbol{x}}_{k+1+m\cdot(l-2)}^{\mathrm{T}}, \hat{\boldsymbol{x}}_{k+m\cdot(l-1)}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{3}$$

The arrangement of these matrices is the following: there are always three matrices of past data, where the past states $\boldsymbol{X}_k^-$, control signals $\boldsymbol{U}_k^-$, and parameters $\boldsymbol{P}_k^-$ form sequence $\boldsymbol{S}_k^-$, the future data is given by the future control signals $\boldsymbol{U}_{k+1}^+$ and parameters $\boldsymbol{P}_{k+1}^+$ forming sequence $\boldsymbol{S}_{k+1}^+$. The prediction target is $\boldsymbol{X}_{k+1}^+$, which also needs to be sampled from the dataset.

### 2.4. Transformer-based Model

In this work, a transformer-based model was employed. While a single Long Short-Term Memory (LSTM) would be sufficient for modeling sequential data, empirical testing demonstrated improved performance for multi-step prediction using a transformer-based architecture that was inspired by the Temporal Fusion Transformer (Lim et al., 2021). This architecture combines LSTMs with transformer-based attention mechanisms. Its strength lies in its ability to handle multiple types of temporal features, such as static covariates, known future inputs, and past observations, while also offering improved interpretability through attention and gating mechanisms. Attention is a mathematical mechanism to detect features and is described best as a mapping of queries and a set of key-value pairs to an output (Vaswani et al., 2017), where the output is computed as a weighted sum of those values. This can then be used as either an encoder or decoder, where the keys, values, and queries come from the same place – the previous layer – but in the decoder, masking is introduced to avoid illegal connections. In an encoder-decoder, the queries come from the previous decoder layer, and the memory keys and values come from the output of the encoder, which is similar to the attention mechanism in sequence-to-sequence models in Gehring et al. (2017), which use LSTM.

To build a transformer model, one typically applies embedding layers to project the inputs into a common hidden dimension suitable for processing by the attention mechanisms. This is often done using fully connected layers. After passing through one or more Multi-Head Attention blocks, followed by feedforward layers, the outputs are projected back to the desired output dimension using another fully connected layer. For a detailed explanation of transformers, the readers are referred to Vaswani et al. (2017).

The architecture of the model used in this work is shown in Figure 4. It consists of a transformer built from two LSTMs as encoder and decoder, a Multi-head attention layer, and a fully connected layer, which makes it a modified version of the Temporal Fusion Transformer from Lim et al. (2021). In this transformer, past data $\boldsymbol{S}_k^-$ is encoded by an LSTM, whose states initialize the decoder processing future data $\boldsymbol{S}_{k+1}^+$. The decoder output passes through a dropout layer (Srivastava et al., 2014), hyperbolic tangent (tanh), temporal self-attention, and a fully connected layer to predict future greenhouse microclimate states $\boldsymbol{X}_{k+1}^+$. The temporal self-attention layer is build from a Multi-head Attention block and dropout applied to its output. The intermediate state $\boldsymbol{m}_t$ and cell state $\boldsymbol{c}_t$ of the encoder are transferred to the decoder.

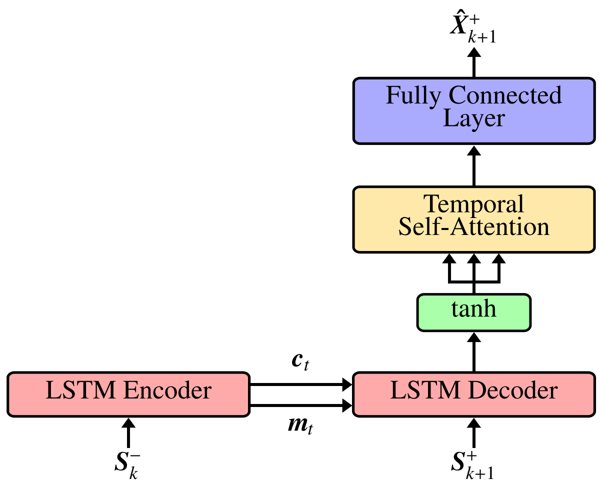


Figure 4: Architecture of the implemented transformer-based model.

### 2.5. Transfer Learning for the Greenhouse Model

The first step in applying transfer learning to the two datasets from the two greenhouses is to compare their state spaces. As analyzed in Section 2.1, both greenhouses have the same number of control signals since all vents were controlled simultaneously. This allows for direct application of transfer learning without modifying the transformer architecture. In this case, the source and target domains are identical because both datasets originate from similar greenhouse systems. However, due to differences in greenhouse structure and dimensions, the microclimate dynamics and their predictions differ slightly from one another. This scenario corresponds to inductive transfer learning (Pan and Yang, 2010).

Although multiple transfer learning strategies exist, preliminary tests indicated that a single approach was sufficient for this study, involving only fine-tuning of the architecture weights and biases. Additionally, Gaussian noise was added to the transformer weights prior to training on the target dataset, following the method proposed in Wu et al. (2022). This initialization perturbs the model parameters away from potential local minima, facilitating further optimization on the target task. Additionally, introducing controlled noise may mitigate catastrophic forgetting (Kirkpatrick et al., 2017) by preserving useful information from the source model while enabling adaptation to the new dataset. This method is defined as

$$\boldsymbol{W}' = \boldsymbol{W} + \mathcal{N}(0, \sigma) \tag{4}$$

$$\boldsymbol{b}' = \boldsymbol{b} + \mathcal{N}(0, \sigma), \tag{5}$$

where $\boldsymbol{W}$ and $\boldsymbol{b}$ are all the weights and biases of the transformer and $\mathcal{N}(0, \sigma)$ is the Gaussian noise with zero mean and standard deviation $\sigma$. This method is later evaluated in Section 3.2.

## 3. Results and Discussions

In the following, both the source and target datasets are used for training, and each is split into a training, validation, and testing dataset via an 80-10-10 split. This means that the source dataset is split into sets of 65, 8, and 8 days, and the target dataset is split into sets of 38, 5, and 5 days of data.

### 3.1. Transformer-based Model Training and Validation

For training the transformer, a sequence with past data $\boldsymbol{S}_k^-$ of length $l = 144$ and stride $m = 20$, covering effectively 24 h, and a sequence with future data $\boldsymbol{S}_{k+1}^+$ of length $l = 72$ and stride $m = 20$, equal to 12 h, were used.

Two prediction models were learned from the source and target datasets by optimizing the transformer models with the Stochastic Gradient Descent (SGD) (Robbins, 1951) optimizer from PyTorch. As before, the data were standardized and divided via an 80-10-10 split into a training, validation, and test set. As it was no random split, the training data consisted of older data than the validation and test set. The hyperparameters were set as follows: A learning rate of $1 \times 10^{-3}$ with a scheduler reducing it to $1 \times 10^{-4}$, a hidden size of 128, an attention head size of 8, one layer for the two LSTMs, and a dropout probability of 0.5. During training, batches of size 16 were used.

For the source dataset, Root Mean Square Error (RMSE) of 1.0038 °C and 4.4707 % for prediction of the inside temperature and humidity were achieved, for the target dataset the respective values are 0.7783 °C and 4.5211 %.

As explained, two different transformers (i.e., for each dataset) were trained to confirm that this type of model is valid to represent microclimate dynamics, regardless of the type of greenhouse.

### 3.2. Transfer Learning Results

The chosen technique of copying the weights and biases and adding Gaussian noise, which was presented in Section 2.5, is evaluated in this section for the transformer-based model trained on the source dataset.

For this type of transfer, a slight alteration of the weights and biases in the source model is expected. Since the number of inputs does not change, the model architecture remains the same. In addition, both datasets were recorded during the same season so all climatic variables have similar minimum and maximum values. Thus, the expected return of this transfer learning approach is the improvement of the predictive capability of the model, mainly to adapt to changes in the underlying dynamics of the greenhouse climate, but without drastically influencing the results.

The transferred model was fine-tuned on the part of the target dataset for training and then tested on the target test dataset. With a standard deviation of the Gaussian noise of 0.01, RMSE of 0.7107 °C and 4.5226 % were obtained for the prediction of the inside temperature and humidity over a 12 h horizon, which is slightly better than the results in Section 3.1. The value of 0.01 for the Gaussian noise was selected by testing different magnitudes for the standard deviation of the noise, ranging from $1 \times 10^{-4}$ to $1 \times 10^{-1}$. It was found, that the maximum Gaussian noise that can be added without worsening the results has a standard deviation of 0.01.

Returning to the motivation for transfer learning in the first place, the question of how much data is necessary for having a working prediction model needs to be answered. Testing is done by varying the size of the training set. As an 80-10-10 split was used, the validation and test sets consist of five days of data each, as there are in total 48 days of data in the target dataset. To have a constant reference frame, the validation and test sets are kept the same for all further tests, while the size of the training set was set to a range of 2 days to 21 days. The minimum is two days, as the selected model needs at least 24 h of past data and 12 h of future data. Therefore, a transformer-based model was trained without and with transfer learning in each training iteration. The same seed initialized both models to prevent the results from being influenced by good initialization.

In Figure 5, it can be observed that, without transfer, RMSE decreases gradually with more data, while with transfer, lower RMSE is achieved after just two days. The slight RMSE increase (top right, with transfer) may stem from adding less relevant data. From these results, it can be concluded that with transfer learning, two days are already enough training data for a reasonably low RMSE and standard deviation. Similar expectations will apply to other transfers when using data from a greenhouse with a different but analogous structure or from another season.

To validate the findings across different two-day periods, the starting point of the training set was varied while keeping its length fixed. RMSE and standard deviation were again computed for comparison, as shown in Figure 6. Without the transfer, the RMSE varies significantly, while it is almost constant with the transfer. The small increase in RMSE for starting days 20 to 25 is most likely related to data that is more different than the other days that are used for training. These results verify again the minimal data required with transfer learning.

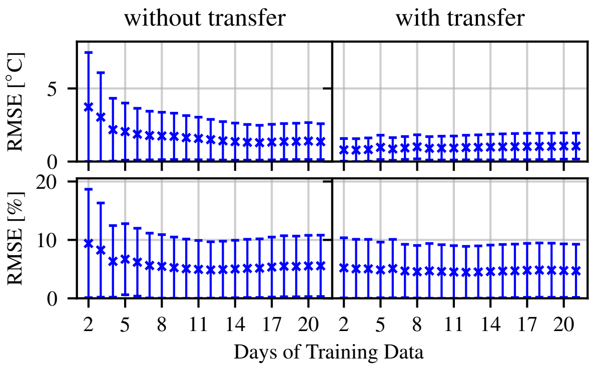


Figure 5: Estimating the minimum training data needed for transfer learning by copying weights and adding noise.

On the other hand, although the RMSE was the selected metric for the quantitative evaluation, visual comparisons also help to analyze the results. To illustrate the impact of transfer learning with minimal data, Figure 7 compares predictions from the models with and without transfer. The predictions were computed by feeding different sequences with 12 h in-between to the model, such that a period of three days was covered. As expected, the transferred model outperforms the non-transferred model very significantly, highlighting its good fit to the real data despite having been transferred using a reduced amount of data.

## 4. Conclusion

This work presented a methodology to model the dynamics of a greenhouse microclimate using a transformer-based model. The model was successfully trained for multi-step prediction of air temperature and relative humidity of a first greenhouse. The transformer was then transferred to a different greenhouse by using transfer learning. According to the results, it can be concluded that the needed amount of training data can be considerably reduced with transfer learning. A model that was transferred and fine-tuned with only two days of additional new training data was able to predict the future states over a 12 h horizon with high accuracy. This approach could be tested for future integration with an Model Predictive Control (MPC) controller. Also, further investigation into different transfers are necessary. This work showed a transfer between similar greenhouses during the same season, while a transfer between different seasons for the same greenhouse would also benefit the reuse of the model. Moreover, a transfer between greenhouses with a different configuration of their vents could be analyzed. For these possible transfers it is necessary to find the minimum amount of additional data from the target datasets and cross-validate the forecasting capabilities, while maintaining the initial assumptions that the amount of data can be reduced and the predictive capability can be increased.

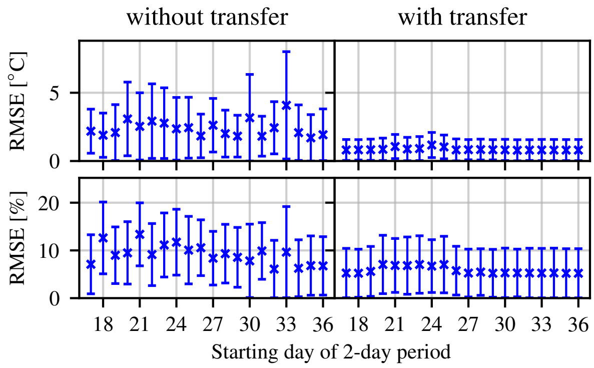


Figure 6: Comparing setups where the training set always consists of two days, but the starting days is changed.

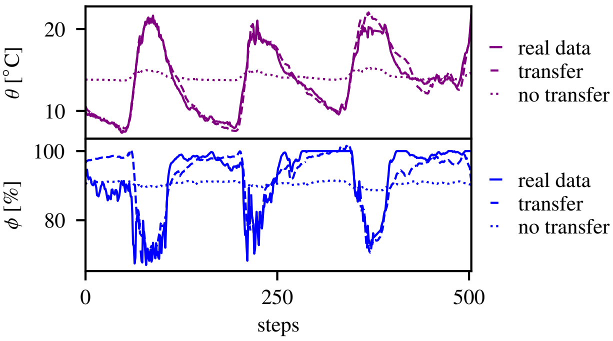


Figure 7: Comparing setups where the training set consists of two days, with and without transfer learning.


## Acknowledgements

This work has been carried out within the framework of the ClimateSmartResearch (CSR) project (HORIZON-CL6-2024-CLIMATE-02-1-101216573), funded by the Horizon Europe Programme (HORIZON Research and Innovation Actions) of the European Union.

This work is a result of the AgroConnect.es facilities used to carry out this research, grant EQC2019-006658-P, funded by MCIN/AEI/10.13039/501100011033 and by ERDF. The authors would like to thank the Cajamar Foundation for providing data from its facilities within the framework of collaboration with the Automatic Control, Robotics and Mechatronics research group of the University of Almería.

The work of Francisco García-Mañas was supported by the Juan de la Cierva Grant (JDC2023-051195-I) of the Spanish State Research Agency.


## References


Fink, M., Daniels, A., García-Mañas, F., Rodríguez, F., Leibold, M., Wollherr, D., 2025. Learning-based model identification for greenhouse climate control. at - Automatisierungstechnik 73 (6), 451–465. DOI: `doi:10.1515/auto-2024-0163`

Gehring, J., Auli, M., Grangier, D., Yarats, D., Dauphin, Y. N., 2017. Convolutional sequence to sequence learning. In: Proceedings of the 34th International Conference on Machine Learning - Volume 70. ICML'17. JMLR.org, p. 1243–1252.

Guo, Z., Feng, L., 2024. Multi-step prediction of greenhouse temperature and humidity based on temporal position attention LSTM. Stochastic Environmental Research and Risk Assessment 38 (12), 4907–4934. DOI: `10.1007/s00477-024-02840-x`

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A., 2018. Transfer learning for time series classification. In: 2018 IEEE International Conference on Big Data (Big Data). IEEE, p. 1367–1376. DOI: `10.1109/bigdata.2018.8621990`

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al., 2017. Overcoming catastrophic forgetting in neural networks. Proceedings of the national academy of sciences 114 (13), 3521–3526.

Lee, J., Im, S., Jeong, J.-S., Lee, T. S., Park, S. H., Shin, C., Ju, H., Kim, H.-J., 2025. Learning hidden relationship between environment and control variables for direct control of automated greenhouse using transformer-based model. Computers and Electronics in Agriculture 235, 110335. DOI: `10.1016/j.compag.2025.110335`

Lim, B., Arık, S. Ö., Loeff, N., Pfister, T., 2021. Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting 37 (4), 1748–1764. DOI: `https://doi.org/10.1016/j.ijforecast.2021.03.012`

Lin, Y.-S., Fang, S.-L., Kang, L., Chen, C.-C., Yao, M.-H., Kuo, B.-J., 2024. Combining recurrent neural network and sigmoid growth models for short-term temperature forecasting and tomato growth prediction in a plastic greenhouse. Horticulturae 10 (3). DOI: `10.3390/horticulturae10030230`

Moon, T., Eek Son, J., 2021. Knowledge transfer for adapting pre-trained deep neural models to predict different greenhouse environments based on a low quantity of data. Computers and Electronics in Agriculture 185, 106136. DOI: `https://doi.org/10.1016/j.compag.2021.106136`

Pan, S. J., Yang, Q., 2010. A survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering 22 (10), 1345–1359. DOI: `10.1109/TKDE.2009.191`

Patil, A., Viquerat, J., Hachem, E., 2023. Autoregressive transformers for data-driven spatiotemporal learning of turbulent flows. APL Machine Learning 1 (4), 046101. DOI: `10.1063/5.0152212`

Robbins, H. E., 1951. A stochastic approximation method. Annals of Mathematical Statistics 22, 400–407.

Rodríguez, F., Berenguel, M., Guzmán, J. L., Ramírez-Arias, A., 2015. Modeling and control of greenhouse crop growth. Springer, Cham, Switzerland. DOI: `10.1007/978-3-319-11134-6`

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15 (56), 1929–1958.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., Polosukhin, I., 2017. Attention is all you need. In: Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems. Vol. 30. Curran Associates, Inc., p. 5998–6008.

Wu, C., Wu, F., Qi, T., Huang, Y., 2022. NoisyTune: A little noise can help you finetune pretrained language models better. In: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Dublin, Ireland, pp. 680–685. DOI: `10.18653/v1/2022.acl-short.76`

Zhao, X., Han, Y., Lewlomphaisarl, U., Wang, H., Hua, J., Wang, X., Kang, M., 2022. Parallel control of greenhouse climate with a transferable prediction model. IEEE Journal of Radio Frequency Identification 6, 857–861. DOI: `10.1109/JRFID.2022.3204363`