



# An Autoencoder Wavelet Based Deep Neural Network with Attention Mechanism for Multistep Prediction of Plant Growth

---

Bashar Alhnaity, Stefanos Kollias, Georgios Leontidis,  
Shouyong Jiang, Bert Schamp and Simon Pearson

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

May 27, 2020

# An autoencoder wavelet based deep neural network with attention mechanism for multistep prediction of plant growth

Bashar Alhnaity<sup>a</sup>, Stefanos Kollias<sup>a</sup>, Georgios Leontidis<sup>a</sup>, Shouyong Jiang<sup>a</sup>, Bert Schamp<sup>b</sup>, Simon Pearson<sup>c</sup>

<sup>a</sup>*School of Computer Science, University of Lincoln, Brayford Pool, Lincoln, UK*

<sup>b</sup>*PCS Ornamental Plant Research, Schaessest raat 18, Dest elbergen, Belgium*

<sup>c</sup>*Lincoln Institute for Agri-Food Technology, University of Lincoln, Riseholme, Lincoln, UK*

---

## Abstract

Multi-step prediction is considered of major significance for time series analysis in many real life problems. Existing methods mainly focus on one-step-ahead forecasting, since multiple step forecasting generally fails due to accumulation of prediction errors. This paper presents a novel approach for predicting plant growth in agriculture, focusing on prediction of plant Stem Diameter Variations (SDV). The proposed approach consists of three main steps. At first, wavelet decomposition is applied to the original data, as to facilitate model fitting to them. Then an encoder-decoder framework is developed using Long Short Term Memory (LSTMs) and used for appropriate feature extraction from the data. Finally, a recurrent neural network including LSTMs and an attention mechanism is proposed for modelling long-term dependencies in the time series data. Experimental results are presented which illustrate the good performance of the proposed approach and that it significantly outperforms the existing models, in terms of error criteria such as RMSE, MAE and MAPE.

*Keywords:* multistep prediction, wavelet analysis, LSTMs, deep neural networks, attention mechanism, time series analysis, plant growth prediction

---

## 1. Introduction

Time-series analysis and prediction has been a research topic of significance in various fields and real-life applications, including smart agriculture and prediction of plant growth, forecasting financial stocks, anomaly, or intrusion, detection, medical imaging and air pollution prediction [10], [1]. Time series data are generally produced as series of observations aggregated in chronological order. Their complexity is generally quite high, which makes their analysis a very challenging task [13]. Due to this nature, using shallow machine learning and neural network models to analyze the data has produced many bottlenecks. As a consequence, the development and use of more complex models, which can automatically extract and learn deep representations from time-series, or image data, has been a topic of major recent work [36, 26, 23, 6].

Recently, Deep Learning (DL) models have produced great progress in agricultural tasks, such as crop management and plant growth. Plants, like other bio-systems, are highly complex and dynamic systems. Modelling plant growth dynamics is a unique challenge, with many different characteristics, e.g., scale of interest, level of description, integration of environmental parameters [1].

This paper proposes a novel deep learning approach for effective prediction of plant growth. It consists of three components: wavelet transformation (WT), encoding-decoding based on LSTM model, and prediction using LSTM with an attention mechanism. WT can assist in smoothing the noise effect existing in time series data. The encoder-decoder part can extract appropriate features from the reconstructed smoothed signal; these features form an appropriate compact representation, on which the final prediction step is based. We propose the use of a model composed of LSTMs blended with an attention mechanism, in order to implement the final prediction of plant growth. As a consequence, the proposed approach is based on the combination of these three methods, and is named as WT-ED-LSTM-AM hereafter. The effectiveness of the WT-ED-LSTM-AM model is validated using real datasets from greenhouses. Moreover, the obtained results are compared with those achieved when using a standard

LSTM network, a multi-layer perceptron (MLP), and a network with gated recurrent units (GRU).

In summary, the main contributions of this paper are the following:

- A novel tree structure for multi-step prediction of SDV. To the best of our knowledge, it is the first time that such an approach is used to predict time series data.
- Extensive experimentation is provided using real-world datasets. The proposed approach is compared with different baseline models; the obtained results show its superiority over these models.

The remainder of this paper is organised as follows: Section 2 reviews related work. Section 3 presents the architecture of the proposed approach, as well as of all utilised models. Section 4 describes in detail the proposed WT-ED-LSTM-AM approach. Section 5 presents the used datasets and the developed experimental study. Finally, Section 6 provides conclusions and future work.

## 2. Related Work

This section provides a short description of existing deep learning prediction models applied to horticulture, and in particular, to plant growth analysis, as well as of attention-based LSTM models for time series analysis and prediction. In general, modelling crop data can be divided into three stages: a) data pre-processing, b) data modelling, c) hybrid analysis approach [45].

Data-driven models (DDM) that are most often used for signal processing include Machine Learning (ML) techniques, Artificial Neural Networks [9], Support Vector Machines [33], and Generalised Linear Models. Those methods have many desirable characteristics, such as: imposing few restrictions and assumptions; ability to approximate nonlinear functions; strong predictive capabilities; flexibility to adapt to multivariate system inputs [3]. According to [40] and [27] Machine Learning, linear polarisation, wavelet-based filtering, vegetation

indices and regression analysis are the most popular techniques used for analyzing agricultural data. Deep Learning (DL) has obtained great popularity in the last few years [17]. DL involves Deep Neural Networks (DNNs), which can generate hierarchical feature structures, leading to more effective representations of the data. This improves learning capabilities, providing higher performance and improving accuracy. A strong advantage of DL is feature learning, i.e., automatic feature extraction from raw data, with features from higher levels of the hierarchy being formed through composition of lower-level features [17]. Consequently, DL can solve more complex problems particularly well [31], provided there exists availability of adequately large data-sets describing the problem. Gonzalez-Sanchez et al. [16] presented a comparative study of ANN, SVR, M5-prime, KNN ML and Multiple Linear Regression for crop yield prediction in ten crop data-sets. In their study, Root Mean Square Error (RMSE), Root Relative Square Error (RRSE), Normalized Mean Absolute Error (MAE) and Correlation Factor (R) were used as accuracy metrics to validate the models. Results showed that M5-Prime achieved the lowest errors across the produced crop yield models. The results of that study ranked the techniques from best to worst, as follows: M5-Prime, kNN, SVR, ANN, MLR. Another study by [4] applied four ML techniques, SVM, Random Forest (RF), Extremely Randomised Trees (ERT) and Deep Learning (DL) to estimate corn yield in Iowa State. Comparison of the validation statistics showed that DL provided the more stable results, overcoming the over-fitting problem. A DL model using LSTM was developed by [1] to predict plant growth, using stem diameter variations as a growth indicator. This paper makes a comparison of the prediction performance of the LSTM model when compared to other baseline models (RF, SVR), showing its improved performance.

Stem diameter is considered a parameter of major importance that describes the growth of plants during vegetative growth stage. The variation of stem diameter has been widely used to derive proxies for plant water status and, is therefore used in optimisation strategies for plant-based irrigation scheduling in a wide range of species. Plant stem diameter variation (SDV) refers to plant

stem periodic shrinkage and recovery movement during the day and night. This periodic variation is related to plant water content and can be used as an indicator of the plant water content change. During active vegetative growth and development, crop plants rely on the carbohydrate gained from photosynthesis and the translocation of photo-assimilates from the site of synthesis to sink organs [46]. The fundamentals of stem diameter variations have been well documented in a substantial amount of literature [44]. It has been documented that SDV is sensitive to water and nutrient conditions and is closely related to the response of crop plants to changes of environmental conditions [22]. Moreover, stem diameter is a parameter that describes the growth of crop plants under abiotic stress during vegetative growth stage. Therefore, it is important to generate stem diameter growth models able to predict the response of SDV to environmental changes and plant growth under different conditions. Many studies emphasize the need to critically review and improve SDV models for assessment of environmental impact on crop growth [20]. SDV daily models have been developed to accurately predict inter-annual variation in annual growth in balsam fir (*Abies balsamea* L). Inclusion of daily data in growth-climate models can improve prediction of the potential growth response to climate by identifying particular climatic events that escape to a classical dendroclimatic approach [12]. However, models for predicting SDV and plant growth using environmental variables have so far remained limited.

Since horticulture management decisions become data-driven, DL is continuously gaining popularity as one of the most successful techniques to model obtained data. As was mentioned above, this paper proposes a new approach for multi-step prediction of plant growth using wavelet transformation (WT), encoder-decoder based on LSTM and RNN-LSTM prediction with an attention mechanism.

### 3. Problem Definition and Components

#### 3.1. Problem definition

The aim of a model for single step time series prediction,  $T$ , is to implement a mapping from a sequence of input data,  $(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_t)$ , to a single output target value,  $y_{t+1}$ :

$$y_{t+1}^{\hat{}} = T(\mathbf{x}_0, \dots, \mathbf{x}_t) \quad (1)$$

where  $\mathbf{x}_t$  is generally an  $M$ -dimensional vector, with elements  $(x_t(0), \dots, x_t(M-1))$  and  $x_t(i), y_t \in \mathbb{R}$ ;  $\hat{y}_t$  is the estimate of the target value  $y_t$ . Eq. 2 shows a multi( $k$ )-step prediction:

$$(\hat{y}_{t+1}, \dots, \hat{y}_{t+k}) = T(\mathbf{x}_0, \dots, \mathbf{x}_t) \quad (2)$$

Model  $T$  is usually estimated through supervised learning, using a collection of training data and respective labels.

#### 3.2. Wavelet transform

The Wavelet transform can be used for data denoising, while handling the non-stationary nature of the collected time series data. In the following we use the wavelet transform for representing, decomposing and reconstructing the original data. Wavelet analysis was firstly introduced by Mallat [30] and since then has been used in a various domains for signal processing [32], image recognition [18], remote sensing data decomposition [34], time series decomposition [11], or medical image analysis and medical diagnosis [41]. The Discrete Wavelet Transform decomposes signals into a low frequency approximation set and several high frequency detailed sets. Thus the original time series, represented as  $X = [\mathbf{x}_0, \dots, \mathbf{x}_{N-1}]$  with  $N = 2^J$ , is transformed as shown in Eqs. 3 and 4 :

$$W_{\varphi}(j, n) = \frac{1}{\sqrt{N}} \sum_k \mathbf{x}_k \varphi_{j,n}(k) \quad (3)$$

$$W_{\psi}(j, n) = \frac{1}{\sqrt{N}} \sum_k \mathbf{x}_k \psi_{j,n}(k) \quad (4)$$

where  $j = 0, 1, \dots, J - 1$ ,  $k = 0, \dots, N - 1$  and  $n = 0, 1, \dots, 2^j - 1$ ;  $\varphi$  and  $\psi$  represent the wavelet function and the scaling function, respectively. To reconstruct the original series, Eq. 5 is used:

$$\mathbf{x}_k = \frac{1}{\sqrt{N}} \sum_n W_\varphi(j, n) \varphi_{j,n}(k) + \sum_{j=j}^{\infty} \sum_n W_\psi(j, n) \psi_{j,n}(k) \quad (5)$$

There are several wavelet families, such as Daubechies (dbN), Coiflets (CoifN) and Symlets (symN). In this paper we use db2 to decompose the original series into one approximation and two detail sets.

### 3.3. Multilayer perceptrons

Multilayer Perceptrons (MLP) [19] have been the main architecture used for supervised learning and classification tasks; they consist of multiple fully connected layers of neurons, with feedforward spread of information. Their training is performed with the backpropagation algorithm.

### 3.4. Long-short term memory

Long short-term memory (LSTM) are one variation of the recurrent neural network (RNN) architecture [21]. They have been able to solve the gradient vanishing problem in long-term time series analysis. The LSTM structure contains three modules: the forget gate, the input gate and the output gate. The forget and input gates control which part of the information should be removed/reserved to the network; the output gate uses the processed information to generate the provided output. LSTMs also include a Cell State, which allows the information to be saved for a long time. The following Eqs. illustrate the feature mapping from inputs of LSTM units:

$$i_t = \sigma(W_i \mathbf{x}_t + U_i h_{t-1} + b_i) \quad (6)$$

$$\tilde{C}_t = \tanh(W_c \mathbf{x}_t + U_c h_{t-1} + b_c) \quad (7)$$



Let us denote by  $f_t$  and  $C_t$  the value of the forget gate and the state of memory cell at time  $t$ , respectively, calculated by:

$$f_t = \sigma(W_f \mathbf{x}_t + U_f h_{t-1} + b_f) \quad (8)$$

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (9)$$

and let  $o_t$  and  $h_t$  denote the values of the output gate and memory cell at time  $t$ , respectively, computed as:

$$o_t = \sigma(W_o \mathbf{x}_t + U_o h_{t-1} + V_o C_t + b_o) \quad (10)$$

$$h_t = o_t * \tanh(C_t) \quad (11)$$

where  $\mathbf{x}_t$  is the input vector to the memory cell at time  $t$ ;  $W_i, W_f, W_c, W_o, U_i, U_f, U_c, U_o$  and  $V_o$  are weight matrices;  $b_i, b_f, b_c$  and  $b_o$  are bias vectors;  $h_t$  is the value of the memory cell at time  $t$ ;  $i_t$  and  $\tilde{C}_t$  are values of the input gate and the candidate state of the memory cell at time  $t$ , respectively.

### 3.5. Gated recurrent units

Gated recurrent units (GRU) are simplified LSTMs [6]. They do not have an output gate, thus there is no control over the memory content. They can be used instead of LSTMs. Further information can be found in [8].

### 3.6. LSTM encoder-decoder models

In LSTM Encoder-Decoder Models, the encoder part compresses the information from the entire input sequence into a vector which is generated from the sequence of the LSTM hidden states. Consequently, the encoder summarizes the whole input sequence into the final cell state vector and passes it to the decoder [5, 29, 43, 28]. The latter uses this representation as initial state to reconstruct the time series, denoted as  $s_t$  in Fig. 1.

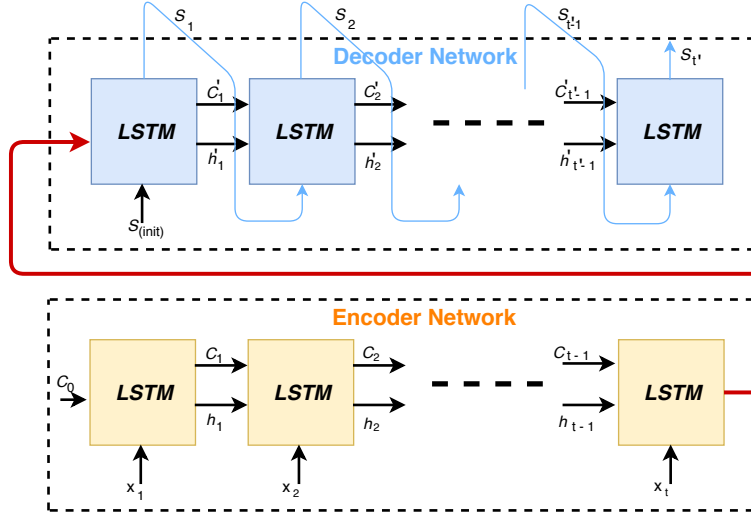


Figure 1: LSTM encoder decoder architecture.

### 3.7. Attention mechanisms

Attention mechanisms have been recently used as a means to improve deep networks' performance, by focusing on feature segments of high significance. They are usually implemented through attentive neural network models [15, 47, 14, 35, 37, 7, 42]. Bahdanau et al [2] introduced an attention mechanism to model a long-term dependence, by generating a context vector as a weighted sum of all provided information. In this paper, the attention mechanism is used both across the different internal LSTM layers, as well as over the LSTM output layers. Prediction of the output signal is derived using the conditional probability distribution of the input signal and of the previous sample of the output signal, i.e.,

$$p(y_i | \mathbf{x}_1, \dots, \mathbf{x}_{i-1}, y_{i-1}) \quad (12)$$

This is, however, impossible to compute in most cases. Eq. 12 is therefore approximated by the non-linear function:

$$g(y_i, h_i, C_i) \quad (13)$$

Where  $g$  is the LSTM,  $h_i$  is the internal state of the LSTM and  $C_i$  is the current context, i.e., a vector holding information of which inputs are important at the current step. The context is derived from both the current state,  $h_i$ , and the input sequence  $\mathbf{x}$ . After the LSTM has stepped through the whole input sequence, the attention mechanism of the network decides on the attention that should be put on the annotation provided at each step. The transition functions of the attentive neural network are described by Eqs. 14-16. The attention mechanism begins by computing  $e_t$ :

$$e_t = v^T \cdot \tanh(W_e \cdot h_t + U_e \cdot d_{t-1} + b) \quad (14)$$

where  $v, b, h_t, d_{t-1} \in \mathbb{R}^n$  and  $W_e, U_e \in \mathbb{R}^{n \times n}$ . The attention score,  $a_t$ , is computed by the softmax function, as follows:

$$a_t = \frac{\exp(e_t)}{\sum_{t=1}^N \exp(e_t)} \quad (15)$$

The context vector,  $C_t$ , is computed as the weighted sum of all internal LSTM states:

$$C_t = \sum_{t=1}^T a_t \cdot h_t \quad (16)$$

## 4. The Proposed Method

### 4.1. Setting up the prediction framework

In the following, we introduce the models and methods, described in the previous Section, in a novel deep prediction framework. The proposed architecture (WT-ED-LSTM-AM) includes wavelet-based transformation of the collected signals, followed by an encoding-decoding step, using LSTM and attention models for final prediction.

Figure 2 shows the proposed approach for SDV multi-step prediction. The target is to predict SDV in the following few hours, based on current information and history sensory signal data.

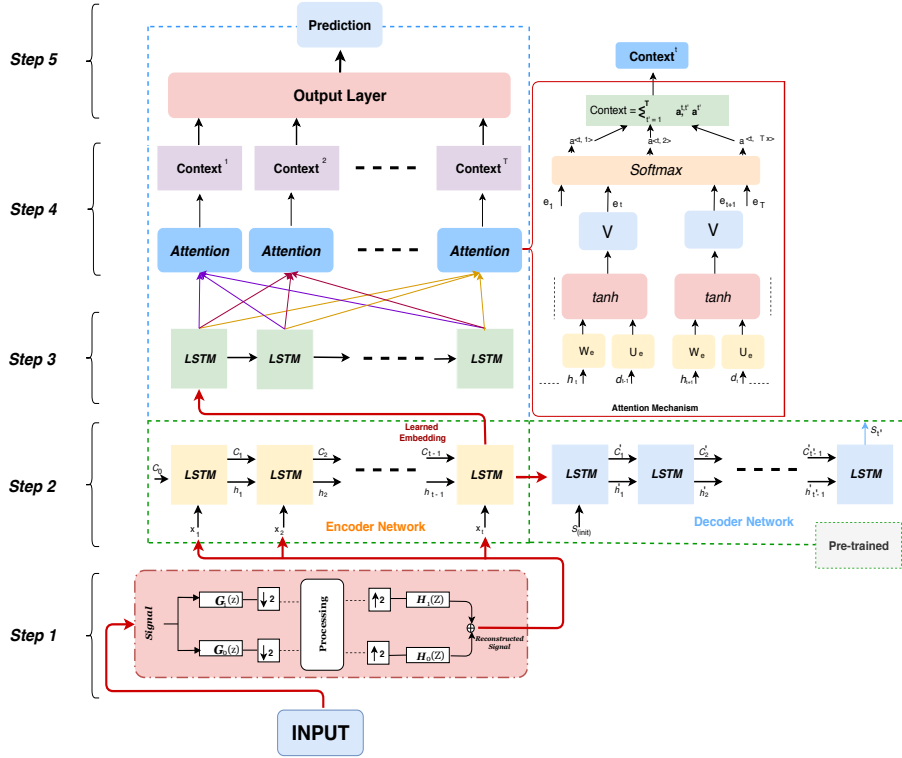


Figure 2: Deep model architecture (WT-ED-LSTM-AM) for SDV prediction.

#### 4.2. The WT-ED-LSTM-AM architecture

As was aforementioned, we propose a novel WT-ED-LSTM-AM architecture for plant growth prediction. This architecture, as shown in Figure 2, is composed of five steps:

- Step 1: Data cleaning and preparation is performed first, by using the wavelet transform (WT). In particular, we decompose each input signal in two components, generating a subsampled (by 2) time series approximation and eliminating noise in the high frequency component. By up-sampling (by 2) and filtering, a reconstructed signal is obtained, which is provided to the next step of our approach.
- Step 2: The encoder-decoder stage is then implemented. The encoder is pre-trained to extract useful and representative embeddings from the

reconstructed time series data ( $x_t$ ); these embeddings can be used next for prediction purposes. Two-layer LSTM cells are used in the encoder implementation. Based on the learned embedding states, the decoder has learnt to generate the (reconstructed) input signal. We designed this step, inspired by the success of video representation learning, where a similar architecture was introduced [43].

- Step 3: The encoder-decoder step consists an intelligent feature extraction component of the proposed approach. Then, an LSTM network, with attention mechanism, is trained to make single, or multi-step prediction, using the learned embedding as input features. LSTMs use the transition functions  $\{h_1, h_2, \dots, h_n\}$  of the embedding states learned in Step 3.
- Step 4: As shown in Figure 2, the attention mechanism is applied to the outputs of each LSTM unit to model a respective long-term dependence. The learned embedding states, the attention weights corresponding to these states, and the respective context, as described in the previous Section, are used for implementing the attention mechanism.
- Step 5: A single layer neural network is responsible for the final prediction of the SDV value, as described in Eq. 18.

$$h_s = \tanh(W_p C + W_x h_n) \quad (17)$$

$$\hat{y} = W_s h_s + b_s \quad (18)$$

## 5. The Experimental Study

An extensive experimental study has been carried out to evaluate the performance of the proposed approach, targeting supervised multi-step prediction of SDV in real-world data sets. The obtained results illustrate the effectiveness and efficiency of the proposed approach in predicting the SDV.

### 5.1. Experimental set up

The proposed architecture was used to predict growth of Ficus plants (*Ficus benjamina*), based on data collected from four cultivation tables in a 90  $m^2$  greenhouse compartment of the Ornamental Plant Research Centre (PCS) in Destelbergen, Belgium. Plant density was approximately 15 pots per  $m^2$ , where every pot contained cuttings.

The experiment started on 23 March 2016. Greenhouse microclimate was set by controlling the window openings, a thermal screen, an air heating system, assimilation light and a CO<sub>2</sub> adding system. Plants were irrigated with an automatic flood irrigation system, controlled by time and radiation sum. Set-points for microclimate and irrigation control were similar to the ones used in commercial greenhouses. The microclimate of the greenhouse was continuously monitored. Photosynthetic active radiation (PAR) and CO<sub>2</sub> concentration were measured with an LI-190 Quantum Sensor (LI-COR, Lincoln, Nebraska, USA) and a carbon dioxide probe (Vaisala CARBOCAP GMP343, Vantaa, Finland), respectively. Temperature and relative humidity were measured with a temperature and relative humidity probe (Campbell Scientific CS215, Logan, UT, USA), which was installed in a ventilated radiation shield.

Stem diameter was continuously monitored on three plants with a linear variable displacement transducer (LVDT, Solartron, Bognor Regis, UK) sensor. The hourly variation rate of stem diameter ( $mm d^{-1}$ ) was calculated as the difference between the current stem diameter and the stem diameter recorded on one hour earlier, at a given time point. Thus, the frequency of collected data has been at one hour basis.

We performed experiments on one-step, two-step and three-step forecasting.

In one-step-ahead forecasting, we used input data collected in previous 15 hours, to predict the SDV value in the current hour.

In two-step-ahead, i.e., 6 hours forecasting, we used input data collected in the previous 6 hours, with a 6-hour stride.

In three-step-ahead, i.e., 12 hours forecasting, we used the previous 12 hours, with a 12-hour stride.

In all experiments, we used the first 70% of data samples as training set, the next 10% of data samples as validation set and the rest 20% of data samples as test set.

### 5.2. Performance evaluation

The Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE) and the Mean Absolute Percentage Error (MAPE) were used to evaluate the performance of prediction models. Formulas of these measures are shown below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left( \frac{A_t - F_t}{A_t} \right)^2} \quad (19)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{|A_t|} \quad (20)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (21)$$

where  $A_t$  denotes the actual values and  $F_t$  the predicted values.

### 5.3. Feature normalization

In all experiments we used min-max normalization (min-max scaling) on the extracted features, re-scaling their values in the range  $[0, 1]$  through the following formula:

$$g_i = \frac{f_i - \min(f)}{\max(f) - \min(f)} \quad (22)$$

where  $g_i$  and  $f_i$  are the values of the normalised and original  $i - th$  feature and  $\min(f)$  and  $\max(f)$  are the minimum and maximum values of the original features.

### 5.4. Experimental results

The experimental results illustrate the very good performance of the proposed methodology, which outperforms respective baseline methods. For comparison purposes, we used the same hyper-parameters in the proposed approach and in the baseline models. A two-layer stacked GRU, LSTM and MLP with

Stochastic Gradient Descent (SGD), learning rate  $l_s = 0.001$  and batch size  $= 32$  were adopted. All models were trained for 100 epochs, using the same training, as well as validation and test data sets. In the proposed method, we used a two layer LSTM encoder-decoder structure, with 128 and 32 neurons respectively. In the prediction model, we used a single layer LSTM with 128 neurons.

The results for the multi-step prediction tasks are shown in Table 1.

Table 1: Performance of the proposed and baseline models for multi-step prediction

Steps	One Step (1hr)			Two steps (6hr)			Three steps (12hr)		
Error Models	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
MLP	0.0034	0.0023	2.72	0.0045	0.0029	2.20	0.0048	0.0027	1.63
GRU	0.0031	0.0022	3.43	0.0039	0.0026	2.74	0.0080	0.0040	1.93
LSTM	0.0031	0.0022	3.27	0.0033	0.0024	2.46	0.0054	0.0031	1.60
Proposed	<b>0.0026</b>	<b>0.0017</b>	<b>2.14</b>	<b>0.0028</b>	<b>0.0021</b>	<b>2.03</b>	<b>0.0029</b>	<b>0.0023</b>	<b>1.35</b>

The performance of the LSTM and GRU models for one-step-ahead prediction were very similar, with the LSTM model showing an (edge) improvement over GRU one, as far the MAPE criterion was concerned. The MLP model performance was lower than LSTM and GRU when considering RMSE and MAE criteria; it scored better than the LSTM and GRU when MAPE criterion was considered.

The proposed approach (WT-ED-LSTM-AM) outperformed all baseline models on all multi-step prediction tasks. Figure 3 illustrates this achievement, over prediction steps ranging from 1 to 12.

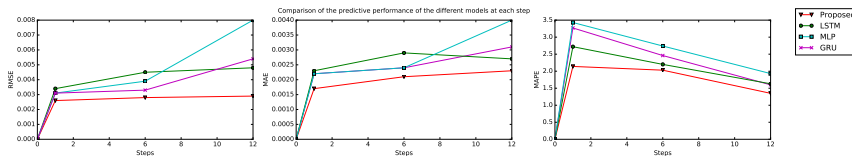


Figure 3: Comparison of the predictive performance of the different models at each step.

Figure 4 shows the accuracy of Ficus growth prediction by all methods for



about 600 data samples. It can be seen that the proposed model successfully performs one-step ahead prediction, outperforming the other methods and providing accurate estimates of almost all peak values in the original data.

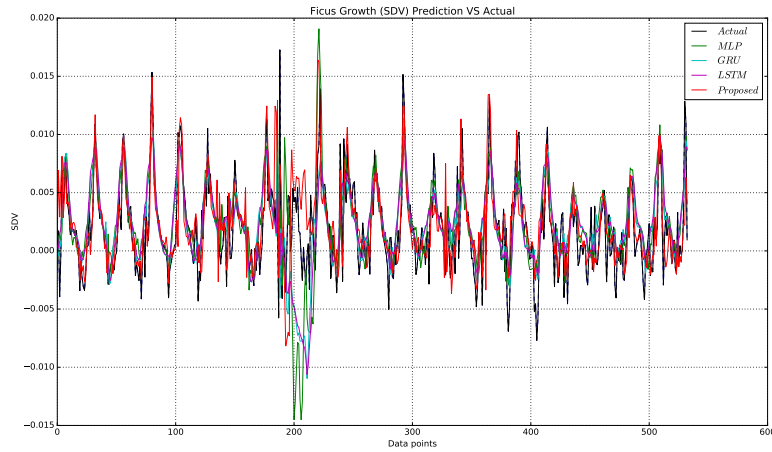


Figure 4: Obtained accuracy in one-step prediction of Ficus growth (SVD) by the proposed and baseline methods

In addition, to compare the distribution of the prediction errors provided by the baseline models with that of the proposed approach, we performed a statistical analysis of them. The histograms of the produced one-step-prediction errors are shown in Figure 5. It can be seen that in the proposed approach, close to 57% of predictions resulted in prediction errors around 0.00 and the remaining 43% prediction errors ranging between -0.010 and 0.015.

The results obtained in the two-step prediction problem are shown in Table 1. The proposed approach outperformed all baseline models, providing lower RMSE, MAE and MAPE values in this case as well. Fig. 6 shows that almost all peak original values are precisely predicted by the proposed approach.

The resulting histograms for two-step-prediction errors are shown in Fig. 7. In the proposed approach, close to 77% of predictions resulted in prediction errors between -0.004 and 0.002; the remaining 23% of prediction errors ranged between -0.008 and 0.008. This greatly outperformed the other baseline models.

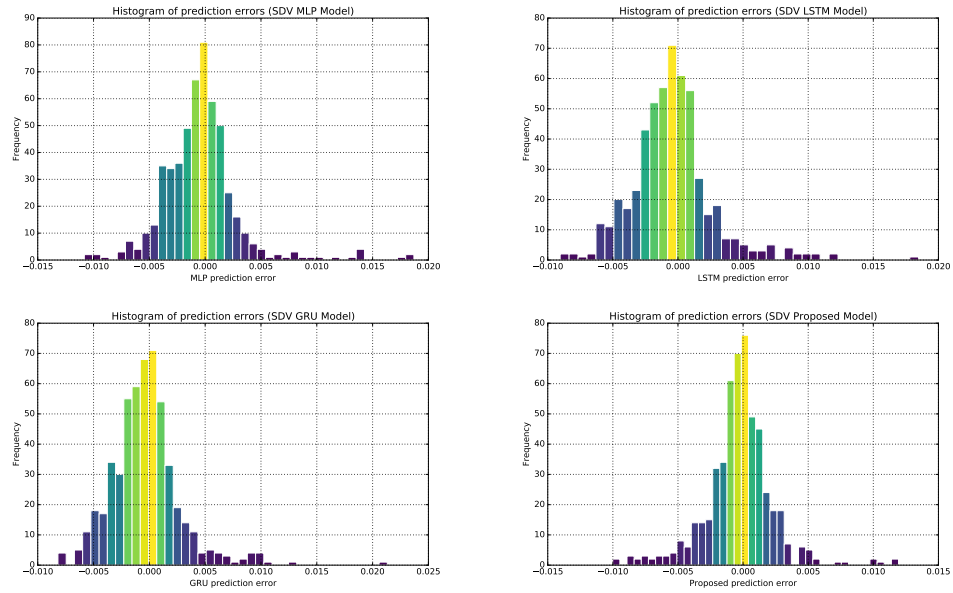


Figure 5: Error distribution for one step prediction (1 hour ahead)

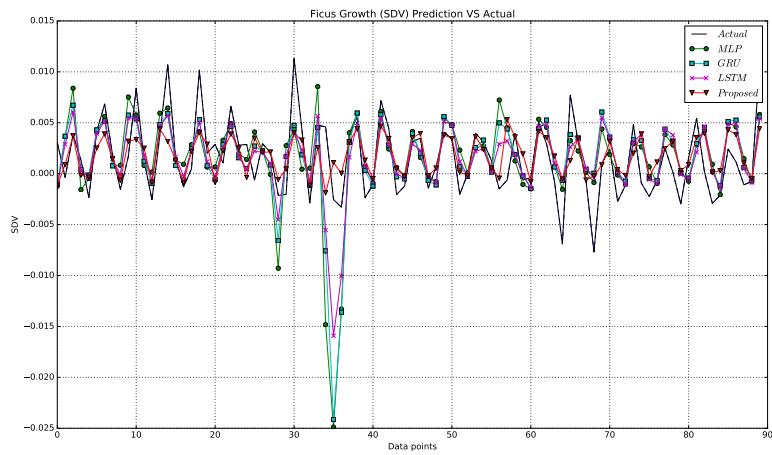


Figure 6: Obtained accuracy in two-step prediction of Ficus growth (SVD) by the proposed and baseline methods

The proposed approach also provided a superior three-step-ahead prediction. Table 1 shows that the proposed approach outperformed the baseline models,

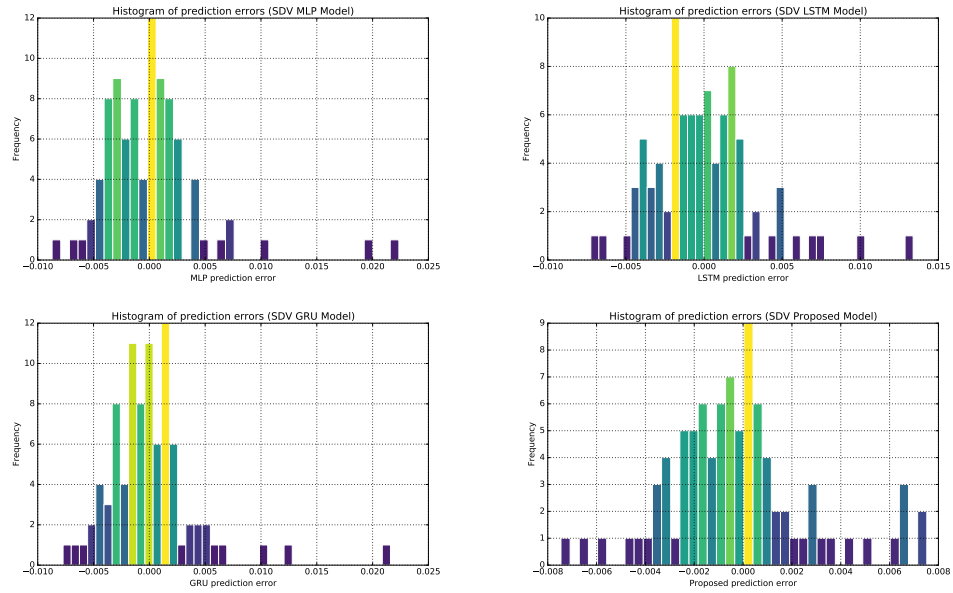


Figure 7: Error distribution for two step prediction (6 hours ahead)

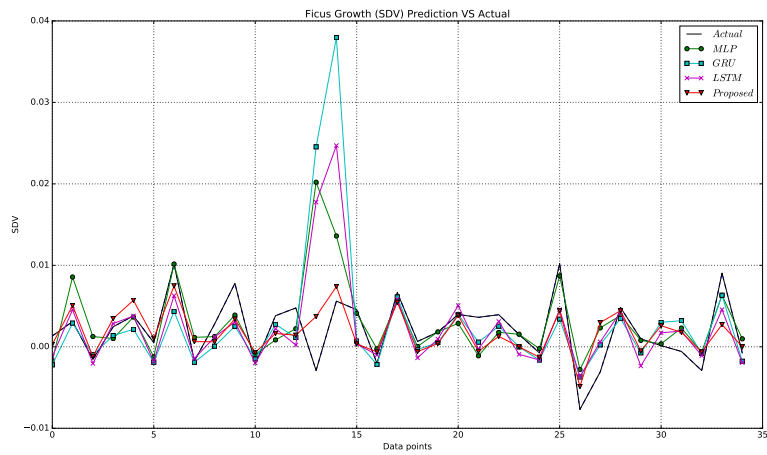


Figure 8: Obtained accuracy in three-step prediction of Ficus growth (SVD) by the proposed and baseline methods

in terms of RMSE, MAE and MAPE criteria. In Fig. 8, it can be seen that all baseline models failed to capture the peak at data sample 14, with the pro-

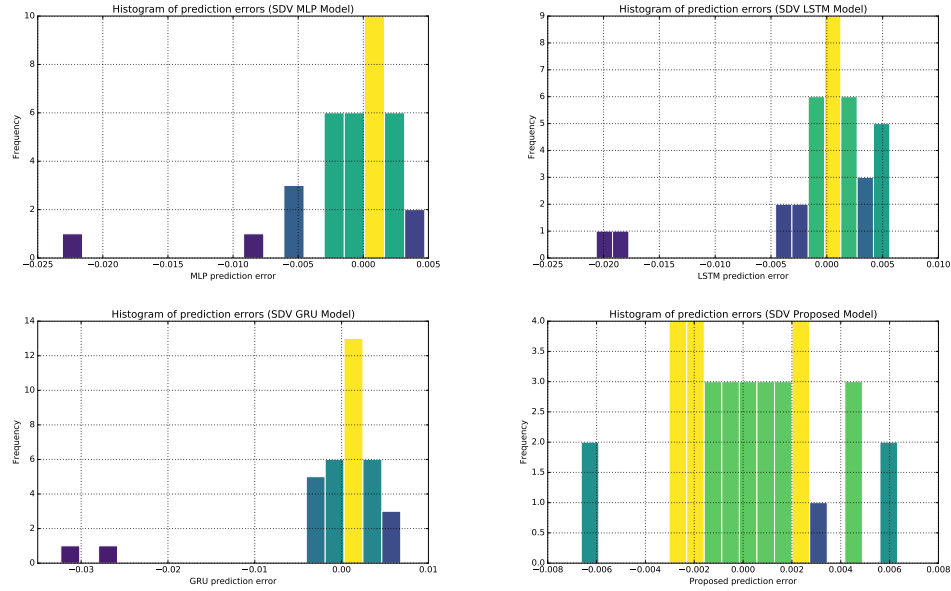


Figure 9: Error distribution for three step prediction (12 hour ahead)

posed approach providing much better estimates of the targeted values than the baseline models. Fig. 9 shows the prediction error distributions for all baseline models, as well for the proposed approach. It shows that close to 56% of predictions provided by the proposed approach, produced prediction errors between -0.002 and 0.002; the remaining 44% of prediction errors ranged between -0.006 and 0.006. In this case, as well, the proposed method outperformed all other baseline models.

## 6. Conclusions and Future Work

This paper proposed a novel multi-step-ahead time series prediction approach. The first step of the proposed method has been to use a wavelet transform to decompose and smooth the original data. As a consequence, a better model fitting could be achieved on the reconstructed signals. The second step introduced an encoder-decoder framework based on LSTMs, which managed to effectively produce appropriate features for multi-step prediction. The

third step which used LSTMs coupled with an attention mechanism was able to successfully implement the prediction tasks.

The proposed approach was used for multi-step-ahead prediction of Ficus Benjamina stem diameter variations, providing high prediction accuracy.

Real-world data have been collected and constituted a data set that was used to evaluate the proposed methodology.

A comparison was carried out, over these real-world data, with state-of-the-art baseline models, showing that the developed approach provides much better prediction results.

Hourly time intervals were used in the input data, as well in our multi-step-ahead predictions. These intervals can be different, i.e., shorter (in minutes, or in seconds), or longer (daily, weekly). Such a scaling of the input data and of the produced predictions require further model developments, as well as experimentation and constitute topics of our future work.

Moreover, another topic of future research is to merge the data driven approach presented in this paper with knowledge-based one, especially for modelling the context, i.e., the relations among the considered variables; we will be adapting former research of ours in fuzzy analysis [38, 39] and description logics [24, 25].

## **Acknowledgements**

This work was done in the framework and with the support of the EU Interreg SMARTGREEN project (2017-2021). We would like to thank all growers (UK and EU), for providing us with the presented data sets. Their valuable feedback, suggestions and comments helped us to improve the quality of this work.

## **References**

- [1] Alhnaity, B., Pearson, S., Leontidis, G., & Kollias, S. (2019). Using deep learning to predict plant growth and yield in greenhouse environments.

*arXiv preprint arXiv:1907.00624*, .

- [2] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, .
- [3] Buhmann, M. D. (2003). *Radial basis functions: theory and implementations* volume 12. Cambridge university press.
- [4] Chlingaryan, A., Sukkariéh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, 151, 61–69.
- [5] Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*, .
- [6] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, .
- [7] Choi, E., Bahadori, M. T., Sun, J., Kulas, J., Schuetz, A., & Stewart, W. (2016). Retain: An interpretable predictive model for healthcare using reverse time attention mechanism. In *Advances in Neural Information Processing Systems* (pp. 3504–3512).
- [8] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, .
- [9] Daniel, J., Andrés, P.-U., Héctor, S., Miguel, B., Marco, T. et al. (2008). A survey of artificial neural network-based modeling in agroecology. In *Soft Computing applications in industry* (pp. 247–269). Springer.

- [10] De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, *22*, 443–473.
- [11] Du, Z., Qin, M., Zhang, F., & Liu, R. (2018). Multistep-ahead forecasting of chlorophyll a using a wavelet nonlinear autoregressive network. *Knowledge-Based Systems*, *160*, 61–70.
- [12] Duchesne, L., & Houle, D. (2011). Modelling day-to-day stem diameter variation and annual growth of balsam fir (*abies balsamea* (l.) mill.) from daily climate. *Forest Ecology and Management*, *262*, 863–872.
- [13] Fu, T.-c. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, *24*, 164–181.
- [14] Gangopadhyay, T., Tan, S. Y., Huang, G., & Sarkar, S. (2018). Temporal attention and stacked lstms for multivariate time series prediction, .
- [15] Geng, Z., Chen, G., Han, Y., Lu, G., & Li, F. (2020). Semantic relation extraction using sequential and tree-structured lstm with attention. *Information Sciences*, *509*, 183–192.
- [16] González-Chávez, M. d. C. A., Carrillo-González, R., Cuellar-Sánchez, A., Delgado-Alvarado, A., Suárez-Espinosa, J., Ríos-Leal, E., Solís-Domínguez, F. A., & Maldonado-Mendoza, I. E. (2019). Phytoremediation assisted by mycorrhizal fungi of a mexican defunct lead-acid battery recycling site. *Science of the Total Environment*, *650*, 3134–3144.
- [17] Goodfellow, I. (2016). Nips 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*, .
- [18] Gupta, D., & Choubey, S. (2015). Discrete wavelet transform for image processing. *International Journal of Emerging Technology and Advanced Engineering*, *4*, 598–602.
- [19] Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.

- [20] Hinckley, T. M., & Bruckerhoff, D. N. (1975). The effects of drought on water relations and stem shrinkage of quercus alba. *Canadian Journal of Botany*, 53, 62–72.
- [21] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9, 1735–1780.
- [22] Kanai, S., Adu-Gymfi, J., Lei, K., Ito, J., Ohkura, K., Moghaieb, R. E., El-Shemy, H., Mohapatra, R., Mohapatra, P. K., Saneoka, H. et al. (2008). N-deficiency damps out circadian rhythmic changes of stem diameter dynamics in tomato plant. *Plant science*, 174, 183–191.
- [23] Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3128–3137).
- [24] Kollia, I., Glimm, B., & Horrocks, I. (2011). Answering queries over owl ontologies with sparql, .
- [25] Kollias, D., Marandianos, G., Raouzaïou, A., & Stafylopatis, A.-G. (2015). Interweaving deep learning and semantic techniques for emotion analysis in human-machine interaction. In *2015 10th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)* (pp. 1–6). IEEE.
- [26] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
- [27] Liakos, K., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18, 2674.
- [28] Lyu, X., Hüser, M., Hyland, S. L., Zerveas, G., & Rätsch, G. (2018). Improving clinical predictions through unsupervised time series representation learning. *arXiv preprint arXiv:1812.00490*, .



- [29] Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). Lstm-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148*, .
- [30] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (pp. 674–693).
- [31] Pan, W. (2016). A survey of transfer learning for collaborative recommendation with auxiliary data. *Neurocomputing*, 177, 447–453.
- [32] Pasti, L., Walczak, B., Massart, D., & Reschiglian, P. (1999). Optimization of signal denoising in discrete wavelet transform. *Chemometrics and intelligent laboratory systems*, 48, 21–34.
- [33] Pouteau, R., Meyer, J.-Y., Taputuarai, R., & Stoll, B. (2012). Support vector machines to map rare and endangered native plants in pacific islands forests. *Ecological Informatics*, 9, 37–46.
- [34] Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., & Immerzeel, W. W. (2011). Improving daily rainfall estimation from ndvi using a wavelet transform. *Environmental Modelling & Software*, 26, 201–209.
- [35] Raffel, C., & Ellis, D. P. (2015). Feed-forward networks with attention can solve some long-term memory problems. *arXiv preprint arXiv:1512.08756*, .
- [36] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85–117.
- [37] Shen, T., Zhou, T., Long, G., Jiang, J., & Zhang, C. (2018). Bi-directional block self-attention for fast and memory-efficient sequence modeling. *arXiv preprint arXiv:1804.00857*, .
- [38] Simou, N., Athanasiadis, T., Stoilos, G., & Kollias, S. D. (2008). Image indexing and retrieval using expressive fuzzy description logics. *Signal, Image and Video Processing*, 2, 321–335.

- [39] Simou, N., & Kollias, S. (2007). Fire : A fuzzy reasoning engine for imprecise knowledge.
- [40] Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in plant science*, *21*, 110–124.
- [41] Singh, R., & Khare, A. (2014). Fusion of multimodal medical images using daubechies complex wavelet transform—a multiresolution approach. *Information Fusion*, *19*, 49–60.
- [42] Song, H., Rajan, D., Thiagarajan, J. J., & Spanias, A. (2018). Attend and diagnose: Clinical time series analysis using attention models. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- [43] Srivastava, N., Mansimov, E., & Salakhudinov, R. (2015). Unsupervised learning of video representations using lstms. In *International conference on machine learning* (pp. 843–852).
- [44] Vandegehuchte, M. W., Guyot, A., Hubau, M., De Groote, S. R., De Baerdemaeker, N. J., Hayes, M., Welti, N., Lovelock, C. E., Lockington, D. A., & Steppe, K. (2014). Long-term versus daily stem diameter variation in co-occurring mangrove species: Environmental versus ecophysiological drivers. *Agricultural and Forest Meteorology*, *192*, 51–58.
- [45] Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming—a review. *Agricultural Systems*, *153*, 69–80.
- [46] Yu, S.-M., Lo, S.-F., & Ho, T.-H. D. (2015). Source–sink communication: regulated by hormone, nutrient, and stress cross-signaling. *Trends in plant science*, *20*, 844–857.
- [47] Yuan, W., Wang, H., Yu, X., Liu, N., & Li, Z. (2020). Attention-based context-aware sequential recommendation model. *Information Sciences*, *510*, 122–134.