Dynamic Bayesian Network for Crop Growth Prediction in Greenhouses

A. Kocian[a,b,*], D. Massa[c], S. Cannazzaro[c], L. Incrocci[b], S. Di Lonardo[d], P. Milazzo[a], S. Chessa[a]

[a]*Department of Computer Science, Pisa University, Italy*
[b]*Department of Agriculture, Food and Evironment, Pisa University, Italy*
[c]*Research Centre for Vegetable and Ornamental Crops Council for Agricultural Research and Economics, Pescia, Italy*
[d]*Research Institute on Terrestrial Ecosystems-National Research Council (IRET-CNR), Firenze, Italy*

Abstract

The paper presents an Internet-of-Things based agricultural decision support system for crop growth. A dynamic Bayesian network (DBN) relates indicative parameters of crop development to environmental control parameters via unobserved (hidden) Markov states. The expectation-maximization algorithm is used to track the states and to learn the parameters of the DBN. The steady state information is then used to derive a predictor for the measurement data a few days ahead. The proposed DBN avoids time-consuming training cultivation cycles, as only data of the current cultivation cycle are available to the algorithm.

Three cultivation cycles of lettuce have been used to test the performance of the proposed DBN. The environmental parameters were temperature, solar irradiance and vapor-pressure deficit. The measurement data include evapotranspiration at granularity equal one day, and leaf-area index and dry weight, at granularity equal one week. It turned out that accurate measurement data prediction a few days ahead is possible even if the number of data samples is low.

Keywords: Internet of Things, Evapotranspiration, Leaf-area, EM Algorithm, Prediction

[*]Corresponding author
Email address: `kocian@di.unipi.it` (A. Kocian)

1. Introduction

1.1. Motivation

The Internet of Things (IoT) technologies Gomez et al. (2019) are gaining broad acceptance and application in many fields: from health-care to enterprise 4.0 to Smart Cities. Even in agriculture these technologies can find numerous applications, both in open field cultivation and in greenhouses, and they are becoming central in the development of precision agriculture, a paradigm that is now increasingly consolidated. The reason of this success lies in the combination of sensing, actuation and communication capabilities of IoT devices, that make them a flexible and adaptable tool to collect data on key crop parameters (as those concerning the environment, the soil and the plant life and growth) and to implement the governing policies on the various plants feeding and assisting the cultures. This approach results particularly effective in greenhouses, where the conditions of the environment, of the root zone and the canopy can be monitored and controlled with high precision. In a typical setting of a technological greenhouse, a number of IoT devices embedding sensors monitor environmental parameters (e.g., air temperature and humidity and solar irradiance), soil/substrate parameters (e.g., moisture, pH, electrical conductivity, ionic concentration) and parameters related to plant physiology and growth e.g., leaf-area index (LAI), accumulated dry weight (DW) and evapo-transpiration (ET). In turn, the fusion of these data through cloud technologies allows the implementation of agricultural decision support systems that forecast growth of cultivation, the development of diseases, performance, etc. The use of such predictive models in real time allows a timely and effective action on crops and the optimization of the resources used, with obvious benefits of sustainability, cost-effectiveness and productivity. In the scientific literature, some of the main models that have been developed aim to make more efficient use of water resources, optimizing irrigation procedures and water management (Incrocci et al., 2019; Massa et al., 2011) while other approaches focus on using optical tools to monitor crop growth and optimize production (Senthilnath et al., 2016b; Padilla et al., 2017).

1.2. Contribution

Our work follows this trend of interdisciplinary research among agriculture, computer science and artificial intelligence. We propose a novel statis-

tical model that relates the indicative parameters of crop development LAI, DW, and ET to the environmental parameters temperature, solar irradiance and VPD at a daily basis. The data processed by the model and the results produced then have the form of a time series of numerically quantifiable measurements. The proposed model is based on a dynamic Bayesian network approach, in which variables influence each other following discrete evolution steps. The model has been developed with the aim to achieve high quality tracking and predictions in an IoT controlled greenhouse Burchi et al. (2018) using a limited amount of data, since some of the data key to the development of the model (like the leaf density) can only be obtained by manual inspection of the crops and hence, are not available in a large quantity. Consequently, our model can be taught based on a few cycles of cultivation and can also make predictions based on the measurement of input data for a few days ahead, with great advantages from the point of view of flexibility and usability in real-world contexts. In this work, we will demonstrate using lettuce that the measurement data can be predicted with high accuracy a few days ahead even if only a few samples of sensor data are available to the algorithm.

The paper is organized as follows. Background and related work are outlined in Section 2. Section 3 presents the system model, followed by a concise description of the Expectation-Maximization Algorithm that is used to track the plant growth parameters, and to predict their evolution. The measurement set-up and the description of the measurement data are outlined in Section 4. The experimental results are described and discussed in Section 5, followed by conclusions and future work in Section 6.

2. Background and Related Work

Leaf area index (LAI) is defined as the total green leaf area per unit horizontal ground surface (dimensionless) and is related to the amount of light that can be intercepted by plants to perform photosynthesis. Hence, LAI is an important parameter to predict plant growth and biomass accumulation through the process of carbon dioxide photoassimilation. Yet, LAI is the most important driving variable for crop transpiration since leaves are the main site for gaseous exchanges between plant and atmosphere. LAI values for various crops differ widely. There are two common approaches to determine LAI and DW: directly i) by taking a statistically significant sample of foliage from a plant canopy. This method is accurate but often tedious and

time-consuming if a teal crop-representative sample would be collected, and ii) by optical sensors. This method is much faster but provides a lower bound on LAI only, as it does not account for leaves that lie on each other. We will build our growth model on the latter method as an indirect measurement of LAI and DW.

Evapotranspiration (ET) is a combined process of evaporation of water from soil or substrate and leaf transpiration of water through the plant tissues to maintain the required crop growth rate and physiological activities. Therefore, ET is closely related to the growth state of the horticultural crop and a primary process affecting irrigation requirements of the crops, to maintain the required crop growth rate. A standard approach to compute the ET [kg m^{-2} d^{-1}] is to apply the Penman-Monteith method Allen et al. (1998), combining energy balance with the mass transfer method. In greenhouse horticulture, the average ET can be approximated by a linear combination of the incoming radiation I [MJ m^{-2} d^{-1}], the LAI [m^2 m^{-2}], latent heat of water vaporization λ [MJ kg^{-2}] and the VPD [kPa] Stanghellini (1987); Baille et al. (1994),

$$\text{ET} = a\frac{I}{\lambda} + b \text{ LAI VPD}. \tag{1}$$

The intercepted global radiation I is function of the global radiation (GR) [MJ m^{-2} d^{-1}], the light extinction coefficient of the canopy (k) and LAI as follows:

$$I = \text{GR} \left(1 - exp(-k \text{ LAI})\right) \tag{2}$$

The coefficients a [dimensionless] and b [kg m^{-2} day^{-1} kPa^{-1}], depending on the particular crop, are obtained by multiple linear regression analysis. Our growth model will be built upon ET measurements on daily basis, but the state-of-the art model in (1) will act as benchmark.

Another major variable, defining the state of the crop is its dry weight (DW). For leaf vegetables, the DW is the weight of the dried shoot that would mostly correspond to produce yield. To determine shoot DW, most of the literature focuses on destructive measurement. Another approach is to model the evolution of DW. Van Henten shows in Van Henten (1994) that the structural DW follows the first order dynamic model

$$\frac{d\text{DW}}{dt} = r \text{ DW} \tag{3}$$

with the constant $r \in \mathbb{R}$. Van Henten's model in (3) will act as benchmark for our growth model that is built on weekly measurements.

The adoption of IoT technologies in agriculture can be observed in the contexts of precision agriculture including protected cultivation Khanna and Kaur (2019); Shi et al. (2019); Zamora-Izquierdo et al. (2019); Muangprathub et al. (2019) but also, more generally, in the agro-industrial and environmental fields Talavera et al. (2017). IoT technologies enable real-time monitoring and control of resources and crops, making it possible to optimize the usage of water, nutrients and energy, and, at the same time, to maximize produce yield and quality. Water management, for example, is a problem for which IoT technologies demonstrated to be particularly well-suited Incrocci et al. (2019); Goap et al. (2018); Du et al. (2017). The use of sensors allows dynamical irrigation policies to be enforced, which minimize the use of water resources while maintaining ideal growth conditions for the crops. A similar argument applies to temperature control in the context of greenhouses Wang and Zhang (2018). As regards the monitoring of plant growth and the maximization of the yield, the application of computer vision and automated digital image analysis technologies has been proposed Easlon and Bloom (2014); Senthilnath et al. (2016a). These technologies make it possible to automatically measure leaf areas and the size/quality of fruits in order to monitor the growth of plants and the production trend.

Obviously, data obtained from a IoT architecture requires to be processed in order to be able to take decisions based on it. Processing involves extracting information from data (e.g. computing growth indicators such as LAI and the evotranspiration rate), but also making predictions on how environmental and soil parameters and growth indicators will evolve in the near future Seginer (2002). Predictions, if sufficiently accurate, allow the farmer to monitor the crop growth better than in real-time, and to quantitatively evaluate (through simulations) the effects of the alternative control actions.

Most of the current predictive models used in agriculture are of physiological-mechanical nature Gary et al. (1998), such as CropSyst and EU-Rotate_N Cilek and Berberoglu (2019). Sensor data is used to calibrate and validate these models. However, they are not able to learn over time. Modern information and communication technology approaches make it possible to generate machines that autonomously learn from data over time Liakos et al. (2018); Balducci et al. (2018); Rehman et al. (2019). In particular, Bayesian methods turned out to be successful in the prediction of aspects of crop growth such as fruit yield Chapman et al. (2018) and disease development Carlson (1970); Bi and Chen (2010). The strength of Bayesian methods is on their ability of making rather accurate prediction even without a huge

amount of available data.

3. A linear dynamic model for crop growth

3.1. Linear Dynamic Model

We develop a dynamic linear model that relates the crop independent environmental data to crop related measurement data. We consider a cultivation of T days with granularity Δ. The study in Carmassi et al. (2007) indicates that the evolution of crop related parameters can be related to environmental data by a dynamic linear model in form of a time series, accommodating the following three kinds of data:

- The control data is organized as K-dimensional column vector $\{\boldsymbol{u}_t : \boldsymbol{u}_t \in \mathbb{R}^K, t \in [1, T]\}$. This data, independent of the crop, is assumed to be deterministic and noiseless.

- The measurement data is organized as column vector $\{\boldsymbol{y}_t : \boldsymbol{y}_t \in \mathbb{R}, t \in [1, T]\}$. This data, depending on the particular crop, is stochastic and noisy.

- The evolution of the hidden states $\{\boldsymbol{z}_t : \boldsymbol{z}_t \in \mathbb{R}^K, t \in [1, T]\}$ relates the control data to the measurement data).

When the system follows a first-order Markov process, the state distribution $p(\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_T)$, abbreviated as $p(\boldsymbol{z}_{1:T})$, has the compact form

$$p(\boldsymbol{z}_{1:T}) = p(\boldsymbol{z}_1) \prod_{t=2}^{T} p(\boldsymbol{z}_t | \boldsymbol{z}_{t-1}). \tag{4}$$

The transition model depends only on the actual state and the previous state. By the Markov property in (4), the joint state-measurement distribution $p(\boldsymbol{z}_{1:T}, \boldsymbol{y}_{1:T})$ has the form

$$p(\boldsymbol{z}_{1:T}, \boldsymbol{y}_{1:T}) = p(\boldsymbol{z}_1) p(\boldsymbol{y}_1 | \boldsymbol{z}_1) \prod_{t=2}^{T} p(\boldsymbol{z}_t | \boldsymbol{z}_{t-1}) p(\boldsymbol{y}_t | \boldsymbol{z}_t). \tag{5}$$

The dynamic Bayesian network (DBN), reflecting the particular factorization of the conditional distributions in (5), is shown in Fig. 1. Each edge corresponds to a conditional dependency, each node corresponds to one of
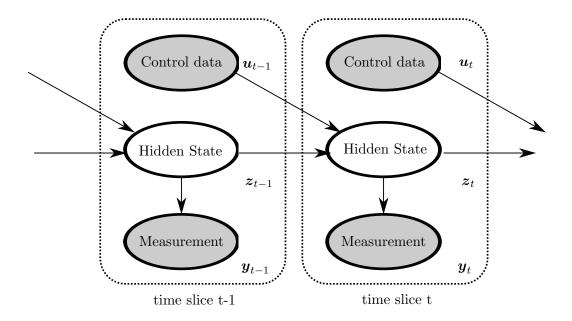
Figure 1: 2-time-slice Bayesian network describing the evolution of the parameters related to the crop.

the three kinds of variables. The arrows indicate the dependencies among the variables. Suppose the DBN describes a linear dynamic model, where variables are all continuous and all of the dependencies are linear Gaussian. Denoting the state matrix $\boldsymbol{A} \in \mathbb{R}^{K \times K}$, the input matrix $\boldsymbol{B} \in \mathbb{R}^{K \times K}$, and the output matrix $\boldsymbol{C} \in \mathbb{R}^{1 \times K}$, the input-output relation of the model is given by

$$\boldsymbol{z}_{t+1} = \boldsymbol{A}\boldsymbol{z}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{n}_t; \qquad \boldsymbol{n}_t \propto \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma_n}) \ i.i.d. \qquad (6)$$

$$\boldsymbol{y}_t = \boldsymbol{C}\boldsymbol{z}_t + \boldsymbol{w}_t; \qquad \boldsymbol{w}_t \propto \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma_w}) \ i.i.d. \qquad (7)$$

$$\boldsymbol{z}_1 = \boldsymbol{\mu}_1 + \boldsymbol{n}_1; \qquad \boldsymbol{n}_1 \propto \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_1) \ i.i.d. \qquad (8)$$

The parameter vector $\boldsymbol{\theta} = \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\Sigma_n}, \boldsymbol{C}, \boldsymbol{\Sigma_w}, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\}$ as well as the state sequence $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T\}$ are unknown and hence, require estimation.

3.2. Tracking of Measurement Data

The EM algorithm is used to iteratively approach the maximum likelihood estimate $\hat{\boldsymbol{\theta}} = \arg\max p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta})$ by postulating the non-observable missing data $\boldsymbol{z}_{0:T}$. Starting from iteration $i = 0$, the E-step of the algorithm infers

the expected log-likelihood $p\left(\boldsymbol{z}_{0:T}, \boldsymbol{y}_{1:T} | \boldsymbol{\theta}\right)$ given the observation $\boldsymbol{y}_{1:T}$ and a guess of the parameter estimate $\boldsymbol{\theta}^{[i]}$ by the M-step:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[i]}) = \mathbb{E}\{\ln p\left(\boldsymbol{z}_{1:T}, \boldsymbol{y}_{1:T} | \boldsymbol{\theta}\right) | \boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\}. \tag{9}$$

The M-step learns from the updated expectations in the E-step, to improve quality of the parameter estimate:

$$\boldsymbol{\theta}^{[i+1]} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[i]}). \tag{10}$$

The sequence of log-likelihood values $\left\{\ln p\left(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}^{[i]}\right)\right\}_{i=0}^{\infty}$ is non-decreasing and converges to a stationary point of $\ln p\left(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}\right)$ Dempster et al. (1977); Wu (1983).

Subsequently we apply above EM algorithm to track the measurement data over a period of T days. The derivation of our tracker, outlined in Appendix A, is an obvious extension of that by Ghahramani and Hinton Ghahramani and Hinton (1996) to linear dynamic models with deterministic control data $\boldsymbol{u}$. After convergence, our tracker at any t outputs the expected state sequence

$$\boldsymbol{z}_t^{[\infty]} \triangleq \mathbb{E}\{\boldsymbol{z}_t|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[\infty]}\} \tag{11}$$

along with error variance

$$\boldsymbol{V}_t^{[\infty]} \triangleq \mathrm{Cov}\{\boldsymbol{z}_t|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[\infty]}\}, \tag{12}$$

Inserting (11) and (12) into (7), the tracked measurement values reads

$$\boldsymbol{y}_t^{[\infty]} = \boldsymbol{C}^{[\infty]}\boldsymbol{z}_t^{[\infty]} \tag{13}$$

along with error variance

$$\Sigma_{\boldsymbol{y},t}^{[\infty]} = \boldsymbol{C}^{[\infty]}\boldsymbol{V}_t^{[\infty]}\left(\boldsymbol{C}^{[\infty]}\right)^T. \tag{14}$$

3.3. Prediction of Measurement Data

The EM algorithm applied to (6)-(8) alternates between prediction and correction by learning from subsequent observation. When the feedback loop is broken, the algorithm may still perform free predictions without response. Following this approach, we design the following q-step ahead predictor. Starting off the state evolution in (6), we take the expectation w.r.t. to

the latest observation and the steady state parameter vector, corresponding to the last estimate by the EM algorithm. Ergo,

$$z_{T+1}^{[\infty]} = A^{[\infty]} z_T^{[\infty]} + B^{[\infty]} u_T. \tag{15}$$

This state estimate has error variance

$$V_{T+1} = A^{[\infty]} V_T \left(A^{[\infty]}\right)^T + \Sigma_n^{[\infty]}. \tag{16}$$

Substituting (15) for (7), we have the one-step ahead measurement predictor

$$y_T^{[\infty]} = C^{[\infty]} z_T^{[\infty]}. \tag{17}$$

By induction hypothesis, it follows for $q > 0$ time steps ahead that

$$y_{T+q}^{[\infty]} = C^{[\infty]} \left(A^{[\infty]} z_{T-1+q}^{[\infty]} + B^{[\infty]} u_T\right), \tag{18}$$

with error variance

$$\Sigma_{y,T+q}^{[\infty]} = C^{[\infty]} \left(A^{[\infty]} V_{T-1+q} \left(A^{[\infty]}\right)^T + \Sigma_n^{[\infty]}\right) \left(C^{[\infty]}\right)^T. \tag{19}$$

Note that the error variance V_{T+1} in (16) depends on the (stochastic) state but is independent of the (deterministic) control data.

3.4. Initialization of the Algorithm

The EM algorithm is sensitive to initialization Hu et al. (2004). Different initial points $\theta^{[0]}$ result in different stationary points of the log-likelihood function. To drive the EM algorithm towards a global maximum of the log-likelihood function, we exploit the structure of available measurement data as follows:

- Since noise is uncorrelated, we may initialize the covariance matrices as
$$\Sigma_n^{[0]} = \Sigma_w^{[0]} = \Sigma_1^{[0]} = \epsilon I, \quad \epsilon \ll 1; \tag{20}$$

- The control matrix $B^{[0]}$ and measurement matrix $C^{[0]}$ are initialized as $B^{[0]} = C^{[0]} = \mathcal{N}(0, I)$;

- Given $\boldsymbol{C}^{[0]}$, the measurement data $\boldsymbol{y}_{1:T}$ and the initial state sequence $\boldsymbol{z}_{1:T}^{[0]}$ are related by

$$\boldsymbol{z}_{1:T}^{[0]} = \left(\boldsymbol{C}^{[0]}\right)^T \left(\boldsymbol{C}^{[0]} \left(\boldsymbol{C}^{[0]}\right)^T\right)^{-1} \boldsymbol{y}_{1:T} \tag{21}$$

according to (7).

- Given $\boldsymbol{B}^{[0]}$, $\boldsymbol{z}_{1:T}^{[0]}$ and $\boldsymbol{y}_{1:T}$, the maximum likelihood estimate of $\boldsymbol{A}$ with Tikhonov regularization has the form

$$\boldsymbol{A}^{[0]} = \left(\boldsymbol{z}_{2:T}^{[0]} \left(\boldsymbol{z}_{1:T-1}^{[0]}\right)^T - \boldsymbol{B}^{[0]}\boldsymbol{u}_{1:T-1} \left(\boldsymbol{z}_{1:T-1}^{[0]}\right)^T\right) \left(\boldsymbol{z}_{1:T-1}^{[0]} \left(\boldsymbol{z}_{1:T-1}^{[0]}\right)^T + \boldsymbol{\Gamma}\right)^{-1}.$$
$$\tag{22}$$

The Tikhonov regularization $\boldsymbol{\Gamma} = \epsilon\boldsymbol{I}$, $\epsilon \ll 1$ enables numerical stability of calculating the matrix inverse.

- Finally, $\boldsymbol{\mu}_1^{[0]} = \boldsymbol{z}_1^{[0]}$.

Note that different choices of $\boldsymbol{B}^{[0]}$ and $\boldsymbol{C}^{[0]}$ may navigate the EM algorithm to different stationary points in the likelihood function.

4. Experimental setting

To illustrate the prediction capability of the EM algorithm, we consider lettuce in a greenhouse at CREA (Research Centre for Vegetable and Ornamental Crops, Council for Agricultural Research and Economics), Pescia, Tuscany, Italy (lat. 43°54′ N, long. 10°42′ E). Plants were transplanted in a soil-less system, consisting in a closed-loop bench (cultivation unit) with drip irrigation and rockwool slabs as substrate, and administered to develop under optimal preserved from abiotic and biotic stresses growing conditions. Plants were watered up to eight times per day. Irrigation started, on average, when 10-20% (v v-1) of the available water was depleted in the growing medium in order to keep moisture quite constant in the root zone. When triggered, irrigation duration was programmed in order to replenish all the nutrient solution present in the slabs with that present in the drainage thank. The nutrient solution absorbed by plants was continuously replenished in the drainage thank by electronic electrovalves when 1/5 of the nutrient solution was absorbed. New fresh nutrient solution was prepared by a fertigation

unit with the following composition in mmol L-1: N-NO3 14.0, N-NH4 1.8, PPO4 1.8, K 8.8, Ca 4.0, Mg 1.0, SSO4 2.2, Cl 0.7, Na 0.8 and microelements according to standard Hoaglands solution. The first growing cycle lasted 35 days ranging from 24 October to 28 November 2018. The second growing cycle lasted 52 days ranging from 21 January to 13 March 2019, and the third growing cycle took 37 days in the period from 4 June - 10 July 2019. The environmental sensors in the greenhouse (Decagon Device Inc., Pullman, WA 99163 USA) measured the environmental control parameters air temperature, air humidity converted into VPD, and global radiation. Their descriptive statistics are reported in Table 1. In contrast, destructive analy-

		Temperature [°C]	Global Radiation [MJ m^{-2} d^{-1}]	VPD [kPA]
Cycle 1	Mean	16.12	3.64	0.54
	Maximum	21.42	7.18	0.88
	Minimum	9.66	0.33	0.13
	Cumulation	N/A	130.18	N/A
Cycle 2	Mean	14.34	5.78	0.78
	Maximum	17.47	9.73	1.25
	Minimum	5.42	0.54	0.12
	Cumulation	N/A	294.9	N/A
Cycle 3	Mean	25.18	17.27	1.15
	Maximum	32.90	19.60	2.81
	Minimum	21.54	12.16	0.72
	Cumulation	N/A	621.66	N/A

Table 1: Descriptive statistics of the environmental control parameters during the cultivation cycles Cycle 1 (24 Oct. - 28 Nov. 2018), Cycle 2 (21 Jan. - 13 Mar. 2019) and Cycle 3 (4 June - 10 July 2019).

ses were carried out to measure LAI ($\Delta = 1$ week), and DW ($\Delta = 1$ week). For the destructive analyses, 15 plants per replicate were collected from three different cultivation units. Finally, ET ($\Delta = 1$ day) was monitored by electronic water meters in three cultivation units different from those used for plant destructive analyses. All data were collected in triplicate. LAI and DW data have been seven times oversampled, to be aligned with the environmental data rate. The maximum number of iterations by the EM algorithm has been set to 100.

For the sake of comparison, the standard ET model in (1) with $a = 0.77$,

$b = 0.077$ [kg m^{-2} day^{-1} kPa^{-1}] and $\lambda = 2.45$ MJ kg^{-2}], acts as benchmark. For the estimation of the intercepted global radiation I we used a k coefficient of 0.66 as suggested by Tei et al. (1996) while LAI was empirically estimated as reported in Carmassi et al. (2013). The model was then calibrated by multiple regression using all the data set. The DW model in (3) has the solution $\mathrm{DW}(t) = \exp(\gamma t + d)$. The constants γ and d have been obtained by exponential regression from the second data set. Following this approach, we obtain $\gamma = 0.106$ [g m^{-2} day^{-1}] and $d = 1.76$ [g m^{-2}].
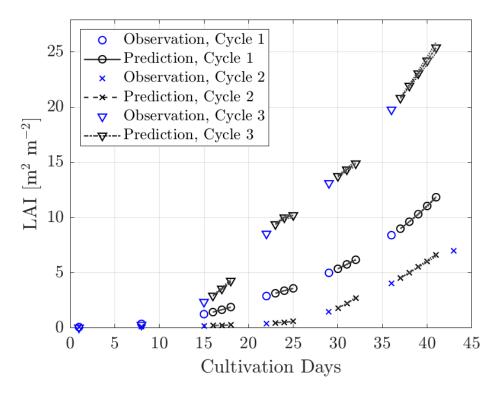
5. Experimental Results and Discussion

Figure 2: Tracking and prediction of the LAI for the lettuce at the growth days $T = \{15, 22, 29, 36\}$ during three cultivation cycles under different environmental conditions: Cycle 1 (late fall), Cycle 2 (late winter), Cycle 3 (early summer).

With above environmental data, the EM algorithm has been used to track the posterior distribution of the measurement data vector until a particular

growing day T, say $T = \{15, 22, 29, 36\}$. At growing day T, the algorithm makes a measurement prediction q days ahead according to (18). Fig. 2 shows the LAI as a function of growing days. The blue curve shows the measured LAI $\boldsymbol{y}_t$, in (7). Clearly, all measurement values beyond T are unknown to the DBN, and are only added to the plot for comparison purposes. The conditional mean value $\boldsymbol{y}_{T+q}^{[\infty]}$ in (18) of our predictor is shown in black. The width of the filled region around the mean value corresponds to twice the error standard deviation $\Sigma_{\boldsymbol{y},T+q}^{[\infty]}$ in (19). It can be seen that the predicted values accurately correspond to the measured values up to a prediction length of $q = 5$ days ahead when the number of data samples is more than four i.e., four weeks. To achieve the same prediction quality for less data samples, the prediction length is about $q = 3$ days ahead. Plants were grown under optimal conditions to avoid any biotic or abiotic stress that could influence the estimated variables. Among other abiotic stress, salinity may influence water and nutrient uptake and plant development. The nominal value for electrical conductivity (EC) of the adopted nutrient solutions was 2.20 dS m^{-1} with a very low concentration of Na and Cl that could possibly accumulate in closed-loop culture (Massa et al., 2011). In the recirculated water, EC averaged 2.32 dS m^{-1} with a coefficient of variability of 7.9% in the different experiments. Lettuce is traditionally considered a species sensitive to salinity in the root zone. However, nutrient and water are much more easily available to plant in hydroponic systems compared with soil cultivation. In many experiments with lettuce grown hydroponically, no variation in dry matter accumulation and plant growth was observed with EC values between 2.0 and 3.0 dS m^{-1} while plant biomass decreased at 1.0 dS m^{-1} (Sago and Shigemura, 2018; da Silva Cuba Cavalho et al., 2018). Yet, a significant decrease in plant water uptake was observed only above EC values of 4.0 dS m^{-1} (Soares et al., 2015).

Fig. 3 reports the cumulative ET for our lettuce. First, we want to point out that the analytical model of Baille in Baille et al. (1994) slightly overestimates the measurement data. This is true for all cycles. For our predictor, it can be seen that with increasing data samples, longer prediction lengths lead to a more accurate forecast. The prediction error averaged over all three cultivation cycles is reported in Fig. 4. When the numbers of cultivation days is equal $T = 15$, a prediction length of $q = 1$ and $q = 5$ results in a prediction error of about 6% and 23%, respectively. This error tends to zero as T becomes large. Clearly, a higher number of cultivation cycles results in a lower error variance and hence, leads to smoother curves. Looking at
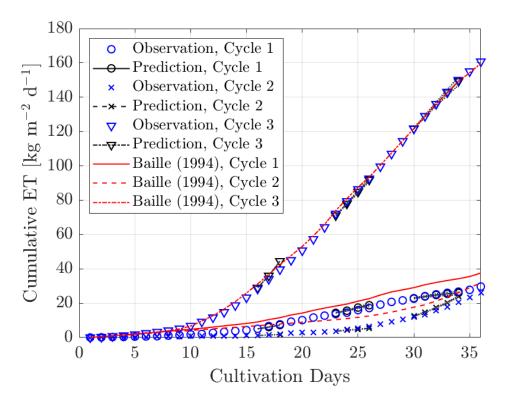
Figure 3: Tracking and prediction of the cumulative evapo-transpiration for the lettuce at the growth days $T = \{15, 22, 29, 36\}$ during three cultivation cycles under different environmental conditions: Cycle 1 (late fall), Cycle 2 (late winter), Cycle 3 (early summer).

the same problem from a different point-of-view, Fig. 5 reports the measured data vs. the predicted data for Cycle 1 (small ET) and Cycle 3 (large ET) with the prediction length as parameter. For the DBN, the algorithm adapts to environmental conditions, and provides ET forecasts within a narrow band around the 1:1 line in an ad-hoc fashion. The Baille model in (1) relies on constant regression coefficients and hence, accurately estimates the measured values during one cycle, namely Cycle 3, but overestimates them during another such as Cycle 1. Generally speaking, the coefficient of determination R^2 of the regression is high for both methods (Cycle 1: $R^2 = 0.997$ (Baille), $R^2 = 1$ (DBN, $q = 1$), $R^2 = 0.998$ (DBN, $q = 5$); Cycle 3: $R^2 = 0.999$ (Baille), $R^2 = 1$ (DBN, $q = 1$), $R^2 = 0.982$ (DBN, $q = 5$)).

Finally, Fig 6 reports the measurement and prediction of dry weight for our lettuce, measured at weekly granularity. It can be seen that the predicted
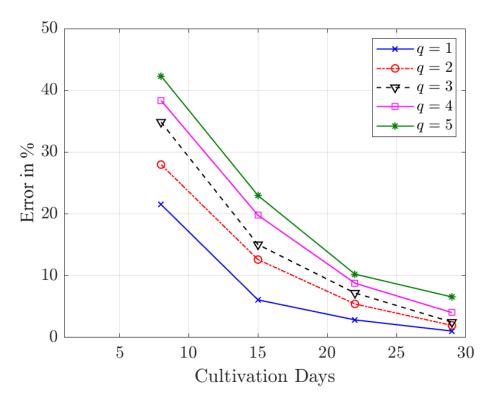
Figure 4: ET prediction error averaged over all three cultivation cycles with prediction length q as parameter.

measurement data is closely related to the true, even if the number of data samples are just a few. When the number of data samples is more than three and four (weeks), a prediction length of, respectively, $q = 3$ and $q = 5$ can be achieved. The behavior is similar to that of predicting LAI in Fig.2. In all cases no historical data from other cycles has been used for prediction. Clearly, the analytical model by Van Henten in (3) is only accurate for the cycle that has been used for calibration.

6. Conclusions

Within the trend of a progressive adoption of IoT and artificial intelligence technologies in agriculture, we considered the specific problem of developing growth models for crop growth in technological greenhouses embedding IoT sensors. The model is based on Bayesian networks and estimates some im-
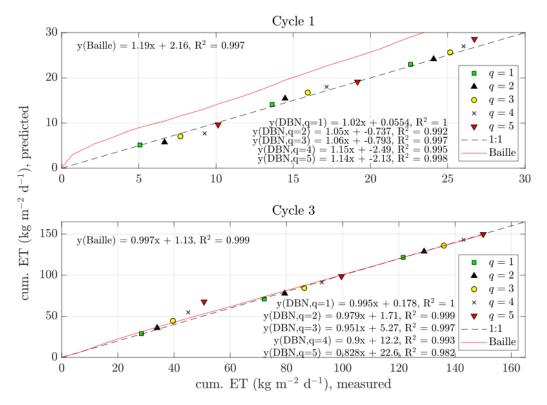
Figure 5: Measured vs. predicted ET for the lettuce with prediction length q as parameter.

portant parameters of crop development (like LAI, DW and ET). It has been tested in an experimental campaign consisting of three cycles of cultivation of lettuce. The experimental results show that the model predicts the ET parameters one day ahead with an error below 6% after 15 cultivation days, and that the values of LAI and DW predicted five days ahead correspond to the measured values.

The analytical models follow a deterministic approach, are simple but depend on a number of environmental and crop specific parameters. Hence, these models are only valid for a specific crop in a confined environment. Our dynamic Bayesian network, in contrast, follows a stochastic approach and hence is valid for a large number of crops with similar statistical properties. The crop related parameters are then learned on-the-fly using IoT based sensor data which are sometimes huge.

Future work includes the experimentation of this model with other crops, and, in perspective, the introduction of additional sensors to monitor other
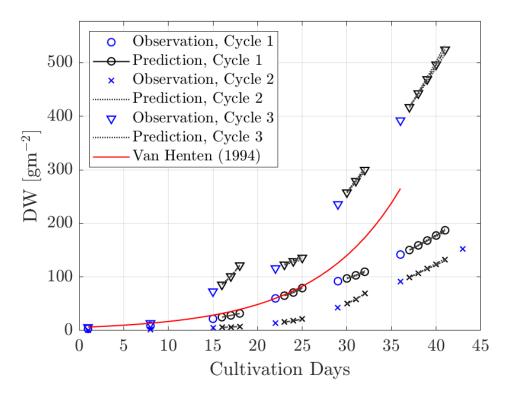
Figure 6: Tracking and prediction of the dry weight for the lettuce at the growth days $T = \{15, 22, 29, 36\}$ during three cultivation cycles under different environmental conditions: Cycle 1 (late fall), Cycle 2 (late winter), Cycle 3 (early summer).

aspects of the plants and soil and their inclusion in the model.

Appendix A. Inference and Learning Algorithm

In this appendix, we derive the EM algorithm for the linear dynamic model in (6)-(8).

By the Markov property in (5), the log-likelihood function of $\boldsymbol{\theta}$ for the

complete data $\mathcal{X}$ is given by

$$\ln p\left(\mathcal{X}|\boldsymbol{\theta}\right) \propto -\ln|\boldsymbol{\Sigma}_1| - (\boldsymbol{z}_1 - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1}(\boldsymbol{z}_1 - \boldsymbol{\mu}_1)$$

$$- T\ln|\boldsymbol{\Sigma}_{\boldsymbol{w}}| - \sum_{t=1}^{T}(\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{z}_t)^T \boldsymbol{\Sigma}_{\boldsymbol{w}}^{-1}(\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{z}_t)$$

$$- (T-1)\ln|\boldsymbol{\Sigma}_{\boldsymbol{n}}|$$

$$- \sum_{t=2}^{T}(\boldsymbol{z}_t - \boldsymbol{A}\boldsymbol{z}_{t-1} - \boldsymbol{B}\boldsymbol{u}_{t-1})^T \boldsymbol{\Sigma}_{\boldsymbol{n}}^{-1}(\boldsymbol{z}_t - \boldsymbol{A}\boldsymbol{z}_{t-1} - \boldsymbol{B}\boldsymbol{u}_{t-1}).$$

$$\text{(A.1)}$$

Let $\boldsymbol{\theta}^{[i]}$ be our parameter estimate at the i-th iteration. Substituting above expression for (9), the E-step of the EM algorithm infers the conditional expectations

$$\boldsymbol{z}_t^{[i]} \triangleq \mathbb{E}\{\boldsymbol{z}_t|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\}$$

$$\boldsymbol{V}_t^{[i]} \triangleq \mathrm{Cov}\{\boldsymbol{z}_t|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\}$$

$$\left(\boldsymbol{z}_t\boldsymbol{z}_t^T\right)^{[i]} \triangleq \mathbb{E}\{\boldsymbol{z}_t\boldsymbol{z}_t^T|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\} = \boldsymbol{V}_t^{[i]} + \boldsymbol{z}_t^{[i]}\left(\boldsymbol{z}_t^{[i]}\right)^T \quad \text{(A.2)}$$

$$\left(\boldsymbol{z}_t\boldsymbol{z}_{t-1}^T\right)^{[i]} \triangleq \mathbb{E}\{\boldsymbol{z}_t\boldsymbol{z}_{t-1}^T|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\} = \boldsymbol{V}_t^{[i]}\boldsymbol{J}_{t-1}^T + \boldsymbol{z}_t^{[i]}\left(\boldsymbol{z}_{t-1}^{[i]}\right)^T,$$

with the short-cut

$$\boldsymbol{J}_{t-1} \triangleq \mathrm{Cov}\{\boldsymbol{z}_{t-1}|\boldsymbol{z}_t, \boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\}\mathrm{Cov}\{\boldsymbol{z}_t|\boldsymbol{y}_{1:T}, \boldsymbol{\theta}^{[i]}\}^{-1}. \quad \text{(A.3)}$$

The computation of (A.2) can be done efficiently by the forward-backward algorithm Rabiner (1989); Minka (1999). Extending the approach in Ghahramani and Hinton (1996) to linear dynamic models with control data, the forward recursion of the forward-backward algorithm yields

$$\boldsymbol{\mu}_t \triangleq \boldsymbol{A}^{[i]}\boldsymbol{\mu}_{t-1} + \boldsymbol{B}^{[i]}\boldsymbol{u}_{t-1} + \boldsymbol{K}_t\left(\boldsymbol{y}_t - \boldsymbol{C}^{[i]}\left(\boldsymbol{A}^{[i]}\boldsymbol{\mu}_{t-1} + \boldsymbol{B}^{[i]}\boldsymbol{u}_{t-1}\right)\right)$$

$$\boldsymbol{V}_t \triangleq \left(\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{C}^{[i]}\right)\left(\boldsymbol{A}^{[i]}\boldsymbol{V}_{t-1}\left(\boldsymbol{A}^{[i]}\right)^T + \boldsymbol{\Sigma}_{\boldsymbol{n}}^{[i]}\right). \quad \text{(A.4)}$$

with the Kalman gain matrix $\boldsymbol{K}$

$$\boldsymbol{K}_t \triangleq \boldsymbol{P}_{t-1}\left(\boldsymbol{C}^{[i]}\right)^T\left[\boldsymbol{C}^{[i]}\boldsymbol{P}_{t-1}\left(\boldsymbol{C}^{[i]}\right)^T + \boldsymbol{\Sigma}_{\boldsymbol{w}}^{[i]}\right]^{-1}, \quad \text{(A.5)}$$

and the short-cut

$$\boldsymbol{P}_t \triangleq \boldsymbol{A}^{[i]} \boldsymbol{V}_t \left(\boldsymbol{A}^{[i]}\right)^T + \boldsymbol{\Sigma}_n^{[i]}. \tag{A.6}$$

The backward recursion reads (A.4)-(A.6) to compute the desired state estimate

$$\boldsymbol{z}_t^{[i]} = \boldsymbol{\mu}_t + \boldsymbol{J}_t \left(\boldsymbol{z}_{t+1}^{[i]} - \boldsymbol{A}^{[i]} \boldsymbol{\mu}_t - \boldsymbol{B}^{[i]} \boldsymbol{u}_t\right),$$
$$\boldsymbol{V}_t^{[i]} = \boldsymbol{V}_t + \boldsymbol{J}_t \left(\boldsymbol{V}_{t+1}^{[i]} - \boldsymbol{P}_t\right) \boldsymbol{J}_t^T, \tag{A.7}$$

where

$$\boldsymbol{J}_t = \boldsymbol{V}_t \left(\boldsymbol{A}^{[i]}\right)^T \boldsymbol{P}_t^{-1}. \tag{A.8}$$

Note that $\boldsymbol{I}$ is the K-dimensional identity matrix. The boundary conditions are

$$\boldsymbol{z}_T^{[i]} = \boldsymbol{\mu}_T$$
$$\boldsymbol{V}_T^{[i]} = \boldsymbol{V}_T. \tag{A.9}$$

The M-step of the EM algorithm in (9) for the linear dynamic model learns the parameter vector $\boldsymbol{\theta}$ by computing the partial derivative of the expected log-likelihood function in (9), setting the result equal zero and solving with respect to the respective parameter. Following this approach, it follows after

straight forward algebraic manipulations

$$\boldsymbol{B}^{[i+1]} = \left(\sum_{t=2}^{T} \left(\boldsymbol{z}_t \boldsymbol{z}_{t-1}^T \right)^{[i]} - \sum_{t'=2}^{T} \boldsymbol{z}_{t'}^{[i]} \boldsymbol{u}_{t'-1}^T \left(\sum_{t=2}^{T} \boldsymbol{z}_{t-1}^{[i]} \boldsymbol{u}_{t-1}^T \right)^{-1} \sum_{t=2}^{T} \left(\boldsymbol{z}_{t-1} \boldsymbol{z}_{t-1}^T \right)^{[i]} \right)$$

$$\times \left(\sum_{t=2}^{T} \boldsymbol{u}_{t-1} \left(\boldsymbol{z}_{t-1}^{[i]} \right)^T - \sum_{t'=2}^{T} \boldsymbol{u}_{t'-1} \boldsymbol{u}_{t'-1}^T \left(\sum_{t=2}^{T} \boldsymbol{z}_{t-1}^{[i]} \boldsymbol{u}_{t-1}^T \right)^{-1} \sum_{t=2}^{T} \left(\boldsymbol{z}_{t-1} \boldsymbol{z}_{t-1}^T \right)^{[i]} \right)^{-1},$$

$$(\text{A.10})$$

$$\boldsymbol{A}^{[i+1]} = \left(\sum_{t=2}^{T} \boldsymbol{z}_t^{[i]} \boldsymbol{u}_{t-1}^T - \boldsymbol{B}^{[i+1]} \boldsymbol{u}_{t-1} \boldsymbol{u}_{t-1}^T \right) \left(\sum_{t=2}^{T} \boldsymbol{z}_{t-1}^{[i]} \boldsymbol{u}_{t-1}^T \right)^{-1}, \qquad (\text{A.11})$$

$$\boldsymbol{\Sigma}_n^{[i+1]} = \frac{1}{T-1} \sum_{t=2}^{T} \left(\boldsymbol{z}_t \boldsymbol{z}_t^T \right)^{[i]} - \left(\boldsymbol{z}_t \boldsymbol{z}_{t-1}^T \right)^{[i]} \left(\boldsymbol{A}^{[i+1]} \right)^T - \boldsymbol{z}_t^{[i]} \boldsymbol{u}_{t-1}^T \left(\boldsymbol{B}^{[i+1]} \right)^T$$

$$- \boldsymbol{A}^{[i+1]} \left(\left(\boldsymbol{z}_{t-1} \boldsymbol{z}_t^T \right)^{[i]} - \left(\boldsymbol{z}_{t-1} \boldsymbol{z}_{t-1}^T \right)^{[i]} \left(\boldsymbol{A}^{[i+1]} \right)^T - \boldsymbol{z}_{t-1}^{[i]} \boldsymbol{u}_{t-1}^T \left(\boldsymbol{B}^{[i+1]} \right)^T \right)$$

$$- \boldsymbol{B}^{[i+1]} \boldsymbol{u}_{t-1} \left((\boldsymbol{z}_t^{[i]})^T - \left(\boldsymbol{z}_{t-1}^{[i]} \right)^T \left(\boldsymbol{A}^{[i+1]} \right)^T - \boldsymbol{u}_{t-1}^T \left(\boldsymbol{B}^{[i+1]} \right)^T \right),$$

$$(\text{A.12})$$

$$\boldsymbol{C}^{[i+1]} = \left(\sum_{t=1}^{T} \boldsymbol{y}_t \left(\boldsymbol{z}_t^{[i]} \right)^T \right) \left(\sum_{t=1}^{T} \left(\boldsymbol{z}_t \boldsymbol{z}_t^T \right)^{[i]} \right)^{-1}. \qquad (\text{A.13})$$

$$\boldsymbol{\Sigma}_w^{[i+1]} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{y}_t \left(\boldsymbol{y}_t^T - \left(\boldsymbol{z}_t^{[i]} \right)^T \left(\boldsymbol{C}^{[i+1]} \right)^T \right)$$

$$- \boldsymbol{C}^{[i+1]} \left(\boldsymbol{z}_t^{[i]} \boldsymbol{y}_t^T - \left(\boldsymbol{z}_t \boldsymbol{z}_t^T \right)^{[i]} \left(\boldsymbol{C}^{[i+1]} \right)^T \right),$$

$$(\text{A.14})$$

$$\boldsymbol{\mu}_1^{[i+1]} = \boldsymbol{z}_1^{[i]}, \qquad (\text{A.15})$$

$$\boldsymbol{\Sigma}_1^{[i+1]} = \left(\boldsymbol{z}_1 \boldsymbol{z}_1^T \right)^{[i]} - \boldsymbol{\mu}_1^{[i+1]} \left(\boldsymbol{\mu}_1^{[i+1]} \right)^T = \boldsymbol{V}_1^{[i]}. \qquad (\text{A.16})$$

This is the derivation of the EM algorithm used to track the states and to learn the parameters of our DBN in Section 3.2.

Declaration of Competing Interest

None.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Food and Agriculture Organization of the United Nations.

Baille, M., Baille, A., Laury, J.C., 1994. A simplified model for predicting evapotranspiration rate of nine ornamental species vs. climate factors and leaf area. Scientia Horticulturae 59, 217–232.

Balducci, F., Impedovo, D., Pirlo, G., 2018. Machine learning applications on agricultural datasets for smart farm enhancement. Machines 6, 38.

Bi, C., Chen, G., 2010. Bayesian networks modeling for crop diseases, in: International Conference on Computer and Computing Technologies in Agriculture, Springer. pp. 312–320.

Burchi, G., Chessa, S., Gambineri, F., Kocian, A., Massa, D., Milano, P., Milazzo, P., Rimediotti, L., Ruggeri, A., 2018. Information technology controlled greenhouse: A system architecture, in: Proc. IoT Vertical and Topical Summit for Agriculture, IEEE, Tuscany, Italy. doi:`10.1109/IOT-TUSCANY.2018.8373044`.

Carlson, G.A., 1970. A decision theoretic approach to crop disease prediction and control. American Journal of Agricultural Economics 52, 216–223.

Carmassi, G., Bacci, l., Bronzini, M., Incrocci, L., Maggini, R., Bellocchi, G., Massa, D., Pardossi, A., 2013. Modelling transpiration of greenhouse gerbera (gerbera jamesonii h.bolus) grown in substrate with saline water in a Mediterranean climate. Scientia Horticulturae 156, 9–18.

Carmassi, G., Incrocci, L., Maggini, R., Malorgio, F., Tognoni, F., Pardossi, A., 2007. An aggregated model for water requirements of greenhouse tomato grown in closed rockwool culture with saline water. Agricultural Water Management 88, 73–82.

Chapman, R., Cook, S., Donough, C., Lim, Y.L., Ho, P.V.V., Lo, K.W., Oberthür, T., 2018. Using Bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. Computers and Electronics in Agriculture 151, 338–348.

Cilek, A., Berberoglu, S., 2019. Biotope conservation in a Mediterranean agricultural land by incorporating crop modelling. Elsevier Ecological Modelling 392, 52–66.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Royal Statist. Soc., Ser. B 39, 1–38.

Du, S., Kang, S., Li, F., Du, T., 2017. Water use efficiency is improved by alternate partial root-zone irrigation of apple in arid northwest china. Agricultural Water Management 179, 184–192.

Easlon, H.M., Bloom, A.J., 2014. Easy leaf area: Automated digital image analysis for rapid and accurate measurement of leaf area. Applications in plant sciences 2.

Gary, C., Jones, J., Tchamitchian, M., 1998. Crop modeling in horticulture: state of the art. Elsevier Scientia Horticulurae 74, 3–20.

Ghahramani, Z., Hinton, G.E., 1996. Parameter Estimation for Linear Dynamical Systems. Technical Report CRG-TR-96-2. University of Toronto.

Goap, A., Sharma, D., Shukla, A., Krishna, C.R., 2018. An iot based smart irrigation management system using machine learning and open source technologies. Computers and electronics in agriculture 155, 41–49.

Gomez, C., Chessa, S., Fleury, A., Roussos, G., Preuveneers, D., 2019. Internet of Things for enabling smart environments: A technology-centric perspective. Journal of Ambient Intelligence and Smart Environments 11, 23–43.

Hu, B., Kocian, A., Piton, R., Hviid, A., Fleury, B.H., Rasmussen, L.K., 2004. Iterative joint channel estimation and interference cancellation using a SISO-SAGE algorithm for coded CDMA, in: Proc. 38th Asilomar Conference on Signals, Systems, and Computers, IEEE, Pacific Grove, CA, USA. pp. 622–626. doi:`10.1109/ACSSC.2004.1399209`.

Incrocci, L., Marzialetti, P., Incrocci, G., Vita, A.D., Balendonck, J., Bibbiani, C., Spagnol, S., Pardossi, A., 2019. Sensor-based management of container nursery crops irrigated with fresh or saline water. Agricultural Water Management 213, 49–61.

Khanna, A., Kaur, S., 2019. Evolution of internet of things (iot) and its significant impact in the field of precision agriculture. Computers and electronics in agriculture 157, 218–231.

Liakos, K., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: A review. Sensors 18, 2674.

Massa, D., Incrocci, L., Maggini, R., Bibbiani, C., Carmassi, G., Malorgio, F., Pardossi, A., 2011. Simulation of crop water and mineral relations in greenhouse soilless culture. Environmental modelling and software 26, 711–722.

Minka, T.P., 1999. From Hidden Markov Models to Linear Dynamical Systems. Technical Report TR-531. MIT Media Lab. MIT, Cambridge, MA, USA.

Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., Nillaor, P., 2019. Iot and agriculture data analysis for smart farm. Computers and electronics in agriculture 156, 467–474.

Padilla, F.M., Peña-Fleitas, M.T., Gallardo, M., Thompson, R.B., 2017. Determination of sufficiency values of canopy reflectance vegetation indices for maximum growth and yield of cucumber. European Journal of Agronomy 84, 1–15.

Rabiner, L.R., 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proc. of the IEEE 77, 257–286.

Rehman, T.U., Mahmud, M.S., Chang, Y.K., Jin, J., Shin, J., 2019. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Computers and Electronics in Agriculture 156, 585–605.

Sago, Y., Shigemura, A., 2018. Quantitative nutrient management reduces nitrate accumulation in hydroponic butterhead lettuces grown under artificial lighting. HortScience 53, 963–967.

Seginer, I., 2002. Sestructures and environment: The penman–monteith evapotranspiration equation as an element in greenhouse ventilation design. Biosystems Engineering 82, 423–439.

Senthilnath, J., Dokania, A., Kandukuri, M., Ramesh, K., Anand, G., Omkar, S., 2016a. Detection of tomatoes using spectral-spatial methods in remotely sensed rgb images captured by uav. Biosystems engineering 146, 16–32.

Senthilnath, J., Dokania, A., Kandukuri, M., Ramesh, K.N., Gautham, A., Omkar, S.N., 2016b. Detection of tomatoes using spectral-spatial methods in remotely sensed RGB images captured by UAV. Biosystems Engineering 146, 16–32.

Shi, X., An, X., Zhao, Q., Liu, H., Xia, L., Sun, X., Guo, Y., 2019. State-of-the-art internet of things in protected agriculture. Sensors 19, 1833.

da Silva Cuba Cavalho, R., Bastos, R.G., Souza, C.F., 2018. Influence of the use of wastewater on nutrient absorption and production of lettuce grown in a hydroponic system. Agricultural Water Management 203, 311–321.

Soares, H., e Silva, Ê.F.F., da Silva, G., Pedrosa, E., Rolim, M., Santos, A., 2015. Lettuce growth and water consumption in nft hydroponic system using brackish water. Revista Brasileira de Engenharia Agrcola e Ambiental 19, 636–642.

Stanghellini, C., 1987. Transpiration of greenhouse crops: an aid to climate management. Ph.D. thesis. Inst. Agricultural Engineering (IMAG), Wageningen, The Netherlands.

Talavera, J.M., Tobón, L.E., Gómez, J.A., Culman, M.A., Aranda, J.M., Parra, D.T., Quiroz, L.A., Hoyos, A., Garreta, L.E., 2017. Review of iot

applications in agro-industrial and environmental fields. Computers and Electronics in Agriculture 142, 283–297.

Tei, F., Aikman, D.P., Scaife, A., 1996. Growth of lettuce, onion and red beet. 2. growth modelling. Annals of botany 78, 645–652.

Van Henten, E., 1994. Validation of a dynamic lettuce growth model for greenhouse climate control. Agricultural Systems 45, 55–72.

Wang, L., Zhang, H., 2018. An adaptive fuzzy hierarchical control for maintaining solar greenhouse temperature. Computers and electronics in agriculture 155, 251–256.

Wu, C.F.J., 1983. On the convergence properties of the EM algorithm. Ann. Stat. 11, 95–103.

Zamora-Izquierdo, M.A., Santa, J., Martínez, J.A., Martínez, V., Skarmeta, A.F., 2019. Smart farming iot platform based on edge and cloud computing. Biosystems engineering 177, 4–17.