Leaf area index soft sensor for tomato crops in greenhouses

F. García-Mañas. F. Rodríguez. M. Berenguel

University of Almería, CIESOL, ceiA3, E04120 Almería, Spain
(e-mail: {francisco.gm; frrodrig; beren}@ual.es)

Abstract: In this work, a soft sensor for tomato crop has been developed using dynamic models to reproduce physical and biological phenomena inside a greenhouse. External weather forecasts were utilised to predict crop growth for a short-term horizon. In addition, data assimilation was performed from observable information of the process, such as measurements captured by sensors, in order to correct uncertainty errors of simulated variables by dynamic models. From the automatic control perspective, the proposed mechanism might allow to implement optimal resource management strategies. Furthermore, crop prediction systems could offer relevant information to farmers as a support tool for decision making.

Keywords: Prediction, virtual sensor, crop monitoring, decision support system, data assimilation.

1. INTRODUCTION

Nowadays, there is a growing concern among international governments to achieve an increasingly efficient agriculture. Despite numerous technological enhancements have been made in recent decades, several policies are arising in order to encourage the implementation of advanced techniques to increase the yield of farms and evolve towards a sustainable bioeconomy model (Egea et al., 2018).

In the coming years, a greater digitalization of agriculture is expected to occur, especially due to the progression of the Internet of Things (IoT) paradigm. IoT technologies might significantly improve the level of existing monitoring in current greenhouses by developing economic sensors to measure variables of interest and simplifying data collection to automate processes or predict future crop scenarios (Tzounis et al., 2017; Woffert et al., 2017; Kochhar and Kumar, 2019). Measurements could be remotely sent to platforms in which Artificial Intelligence and Big Data techniques are applied to optimally analyse and interpret the collected information (Jayaraman et al., 2016; Guirado-Clavijo et al., 2018).

In greenhouses, some relevant variables are yet difficult to be measured by sensors, such as Leaf Area Index (LAI), one of the most important indicators to evaluate crop growth. At present, LAI is commonly measured by manually performing destructive tests to some plants of the crop. This tedious procedure encourages to develop an alternative technique to avoid destructive tests and to offer a method for continuous monitoring. For instance, new artificial vision techniques are aimed to estimate growth based on crop images (Kaiyan et al., 2014). Also, neural networks seem to be useful (Qaddoum et al., 2013; Wang et al., 2017).

In this work, a different proposal is presented, focusing on the concept of soft sensors, also named virtual sensors. A soft sensor has been developed to offer an estimation for LAI by using available information from other real sensors inside a greenhouse. The LAI estimation is generated with a dynamic model specifically developed for tomato crops (Jones et al., 1999). Soft sensors based on dynamics models have remarkable versatility to be improved. New approaches are emerging to combine components from different classical models in order to obtain a better overall performance in tomato crop growth estimation (Kuijpers et al., 2019).

2. MATERIALS AND METHODS

2.1 Description of the greenhouse

In this work, a traditional greenhouse was used to obtain climate and crop data under real production conditions. The “parral-type” greenhouse (Fig. 1) is located at “Las Palmerillas” Experimental Centre (Almería, Spain), property of Cajamar Foundation.

Fig. 1. Traditional “parral-type” greenhouse. Left: exterior view. Right: interior view.

The total surface of the greenhouse is 877 m² (37.8 × 23.2 m). Under the polyethylene cover, an area of 600 m² is dedicated for the crop, which is distributed in rows oriented from north to south, with a longitudinal slope of 1%. Planting is done in coconut fibre bags.

Natural ventilation is automated to control inside air temperature and humidity by means of five zenithal windows (8.36 x 0.73 m) and two lateral windows (32.75 x 1.9 m) situated on the north and south sides of the greenhouse.
2.2 Experimental data

Experimental data for this work came from three different sources: (i) climate variables monitored outside and inside the greenhouse, (ii) manual measurements to determine crop growth, and (iii) weather forecast provided by the Spanish State Meteorological Agency, AEMET.

The greenhouse at “Las Palmerillas” has an advanced system for measuring and controlling climatic variables, developed by the Automatic Control, Robotics and Mechatronics research group (ARM) of the University of Almeria. Outside the greenhouse, a meteorological station registers air temperature and relative humidity, wind velocity, solar radiation, and CO₂ concentration in the air. Inside the greenhouse, distributed sensors measure air temperature and relative humidity, CO₂ concentration in the air, global radiation, photosynthetic active radiation (PAR), and soil surface temperature, among other variables not utilised in this work. All variables are registered every 30 seconds.

Tomato crop growth was measured manually by performing destructive tests applied to some plants randomly selected from the entire crop. In each test, plants were cut and broken down into leaves, stems, and fruits, allowing the area of the leaves to be measured and the fresh weight of each part to be determined. Subsequently, samples were dried for 24-48 hours, and finally weighed to obtain the total dry weight. Destructive planting was carried out every 30 days, approximately. Periodic maintenance tasks of the crop were also registered, since pruning and deleafing also affect leaf area index.

External weather forecasts provided by AEMET were obtained from the high-resolution atmospheric model HARMONIE-AROME (Bengtsson et al., 2017). This powerful model performs data assimilation from satellite observations, ground weather stations, radiosondes and sensors in the fuselage of commercial aircrafts. For this work, AEMET periodically sent weather data to a private server at University of Almeria. Everyday, four climate forecasts were provided at 00:00, 06:00, 12:00 and 18:00, Coordinated Universal Time, UTC. Each forecast had a 48 hours prediction horizon for the following variables with a 15 minutes sampling interval: air temperature, air relative humidity, solar radiation, wind velocity, wind direction and total accumulated rainfall.

2.3 Tomato growth model

In this work, the reduced TOMGRO model (Jones et al., 1999) was utilised to simulate tomato crop growth. This model allows to estimate five state variables, as showed in the schematic representation in Fig. 2. The model consists of a set of equations to calculate crop photosynthesis, respiration, and other metabolic activities. The original model was adapted to Mediterranean greenhouses conditions by including the effects of irrigation and fertilizers (based on electrical conductivity). A more extensive description can be found in (Rodríguez et al., 2015).

Leaf Area Index is a reference variable to indicate crop growth. In the reduced TOMGRO model, LAI evolution in time is determined by a differential equation showed in (1), as a function of other crop variables such as number of nodes, \(X_N\). This equation is also affected by an unitless function \(f_{\text{LAI}}(\bar{X}_{T,a})\), which reduces LAI depending on the daily average of inside greenhouse air temperature, plant density, \(\rho\), maximum leaf area expansion per node, \(\delta_l\), and a projection coefficient related to the number of nodes, \(N_y\).

\[
\frac{dX_{\text{LAI}}}{dt} = \rho \delta_l f_{\text{LAI}}(\bar{X}_{T,a}) e^{(\beta_l(X_N-N_y))} \frac{dX_N}{dt}
\]

(1)

2.4 Greenhouse climate model

The microclimate inside a greenhouse can be simulated by a simplified pseudo-physical dynamic model, completely described by Rodriguez et al. (2015). This model was obtained by simplifications of first principles, such as energy balances (e.g. heat transfer flows) and mass balances (e.g. water vapor in the air). State variables for this model are inside air temperature, inside air relative humidity, and inside solar radiation. In this work, a sub-model was included to simulate inside soil surface temperature as an internal disturbance.

In Fig. 3, a schematic representation of the model is exposed. Red coloured inputs correspond to disturbance variables, fundamentally associated to external weather conditions, and the blue input is a manipulated variable (actuator). Simulated state variables are coloured in green as model outputs.

Inside air temperature, \(X_{T,a}\), is calculated in (2) as a differential equation considering greenhouse inside heat fluxes, \(Q\), and physical constants such as air specific heat coefficient, \(c_{\text{ph,a}}\), air density, \(c_{\text{den,a}}\), greenhouse volume, \(c_{\text{vol,g}}\), and inside soil surface area, \(c_{\text{area,ss}}\). 

\[
c_{\text{ph,a}} \cdot c_{\text{den,a}} \cdot c_{\text{vol,g}} \cdot \frac{dX_{T,a}}{dt} = \sum Q + \sum Q(X_{T,a})
\]

(2)

Fig. 2. Schematic representation of tomato growth model.

Fig. 3. Schematic representation of greenhouse climate model.
Inside air relative humidity, \(X_{HA,ar}\), is determined with the mass balance expressed in (3), where \(M_{trp,cr}\) is the crop transpiration flux, and \(M_{ven,a-e}\) is the outflow corresponding to natural ventilation.

\[
\frac{c_{den,ss}}{c_{area,ss}} \cdot \frac{dX_{HA,a}}{dt} = M_{trp,cr} - M_{ven,a-e}
\]

Solar radiation inside the greenhouse, \(V_{r,x,cr}\), is calculated by the static expression showed in (4), where \(D_{r,x,cr}\) is the external solar radiation, and \(V_{tsw,g}\) is a short wave transmission coefficient depending on the type of greenhouse cover.

\[
V_{r,x,cr} = V_{tsw,g} \cdot D_{r,x,cr}
\]

(4)

Finally, soil surface temperature, \(x_{T,ss}\), can be simulated with equation (5), where \(c_{sph,ss}\) is the specific heat coefficient for soil surface material, \(c_{den,ss}\) is the soil surface material density, and \(c_{th,ss}\) is the thickness of the soil surface.

\[
c_{sph,ss} \cdot c_{den,ss} \cdot c_{th,ss} \cdot \frac{dx_{T,ss}}{dt} = \Sigma Q_{ss}
\]

This equation was implemented as a simplified sub-model to be only executed when no real measurements were available for soil surface temperature. Thus, this sub-model can offer an estimation for this variable as an internal disturbance for greenhouse climate.

### 2.5 Models calibration and validation procedure

A genetic algorithm (Houck et al., 1995) was employed to achieve constants calibration by minimizing the error between simulated dynamics and real dynamics of the process. The utilised error function corresponds to the root of sum of quadratic errors and is presented in the following expression:

\[
Error = \sqrt{n \sum_{k=1}^{n} (x_{real,k} - x_{sim,k})^2}
\]

where \(x_{real,k}\) are real data from process, \(x_{sim,k}\) are simulated data from a dynamic model, and \(n\) is the total number of samples (referred to time instants, \(k\)).

Based on the sensitivity analyses originally performed by Rodriguez (2002), constants with greater impact on simulated dynamics were calibrated. Constrained values (limits) were imposed during the calibration procedure to ensure that calibrated constants retain their physical meaning.

Validation methodology consisted in performing several simulations for each model with real experimental data that offered dynamic variety to graphically evaluate the behaviour of the models under different scenarios. For each validation test, absolute errors between simulated variables and real measurements were calculated. Mean Absolute Error (MAE) was selected as an index to assess the quality of models’ simulations, as it represents the average deviation of simulated variables against real measurements.

### 2.6 Software

Dynamic models’ equations and prediction mechanism functions were programmed in MATLAB® (MathWorks, Massachusetts, USA), version 2018b. In addition, analysis tools such as MATLAB® Profiler were utilised to optimize the execution time of the programmed code. Simulations with real experimental data were also performed in this software.

### 3. RESULTS AND DISCUSSIONS

#### 3.1 Tomato growth model validation

Tomato growth model was validated with real crop data from a complete season between September 2015 and May 2016 (240 days). The model was simulated using the actual values registered for inside air temperature and PAR radiation, as shown in Fig. 4. Electrical conductivity was controlled throughout the season, with an average value of 2.6 dS·m⁻¹. CO₂ concentration in the air inside the greenhouse was not measured continuously. Hence, in order to simulate the model, it was assumed an average value of 380 ppm for CO₂ concentration, as no enrichment systems were utilised.

Fig. 4. Inputs for tomato growth model simulation. Season 2015-2016.

Plant density was established in 1.4 plants·m⁻². During the season, eight destructive tests were performed to determine real values of crop biomass. Additionally, nine pruning tasks were registered. The amount of LAI retired from the plants in each pruning was provided to the model so that simulated LAI could be modified by subtracting the pruned value.

Values for model constants were taken from baseline studies previously published in the literature (Jones et al., 1991, 1999). Some constants were calibrated in previous works to reflect some characteristics for the crop variety, as shown in Table 1.

| Table 1. Calibrated constants for tomato growth model |
|---------------------------------|---------|---------|
| Constant | Value | Units     | Affects to         |
|  \(c_{Nm}\) | 0.56  | nodes·day⁻¹ | Nodes             |
|  \(c_{Np}\) | 4     | nodes     | LAI               |
|  \(c_{P}\) | 0.74  | g·d.w.·g⁻¹CH₂O | Growth rate     |
|  \(c_{ext}\) | 0.65  | -        | Photosynthesis    |
|  \(c_{rm}\) | 0.016 | g·CH₂O·g⁻¹·d.w.·day⁻¹ | Respiration |
|  \(c_{Tcr} \) | 28    | °C       | Fruit dry weight  |
Validation results are presented in Fig. 5, comparing model simulation for state variables against real biomass data. Real data tendency was calculated as a linear interpolation of the scattered samples from destructive tests. Standard deviation of real samples was determined for each destructive test and is represented as error bars. Results reveal a good correlation between the simulated variables and the actual samples, considering that a simplified model has been utilised. It should be highlighted the complexity of validating the model with sporadic biomass samples in a season with several pruning.

Table 2. (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variation interval</th>
<th>MAE</th>
<th>Max.</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total d. w.</td>
<td>[0.42, 1500]</td>
<td>61.59</td>
<td>128.75</td>
<td>34.16</td>
</tr>
<tr>
<td>Fruit d. w.</td>
<td>[0, 1073]</td>
<td>44.23</td>
<td>118.39</td>
<td>35.77</td>
</tr>
<tr>
<td>Mature d. w.</td>
<td>[0, 955.20]</td>
<td>38.18</td>
<td>96.57</td>
<td>35.66</td>
</tr>
</tbody>
</table>

3.2 Greenhouse climate model validation

Greenhouse climate model calibration must be achieved as proposed by Rodríguez (2002). The methodology consists in considering two scenarios: (i) empty greenhouse (without crop), and (ii) greenhouse with crop inside. For this work, calibration and validation for climate model was carried out at the beginning of the current season in August 2019. The greenhouse was initially empty and crop transplantation was performed in the last weeks of August 2019.

In the calibration procedure, five constants (Table 3) were adjusted. For validation test, the climate model was simulated with real input data as shown in Fig. 6. Real inside soil surface temperature was also utilised as an input. Model validation results are presented in Fig. 7, with a good adjustment of the model, since simulated dynamics clearly evolves in concordance with real dynamics, even when rapid changes occur for external solar radiation.

Table 3. Calibrated constants for climate model

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{ven,d}$</td>
<td>Ventilation discharge coefficient</td>
<td>-</td>
</tr>
<tr>
<td>$c_{ven,cv}$</td>
<td>Ventilation wind coefficient</td>
<td>-</td>
</tr>
<tr>
<td>$c_{cvn,ss-a}$</td>
<td>Convection coefficient between soil surface and air</td>
<td>W·m⁻²·K⁻¹</td>
</tr>
<tr>
<td>$c_{cvn-cov,a-e}$</td>
<td>Thermal loss coefficient in the cover between outside and inside air</td>
<td>W·m⁻²·K⁻¹</td>
</tr>
<tr>
<td>$V_{tsw,g}$</td>
<td>Short wave transmission coefficient (cover)</td>
<td>-</td>
</tr>
</tbody>
</table>

In view of the MAE error analysis (see Table 4), all state variables present significantly reduced errors for their range of variation, agreeing with previous studies (Rodríguez et al., 2015). Therefore, it can be concluded that the model successfully reproduces the indoor climate of the greenhouse.

Table 4. Error analysis for climate model validation

<table>
<thead>
<tr>
<th>Variation interval</th>
<th>Temp. (°C)</th>
<th>RH (%)</th>
<th>Rad. (W·m⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[22.55, 42.08]</td>
<td>[29.48, 80.78]</td>
<td>[0, 497.30]</td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>0.96</td>
<td>2.55</td>
<td>16.44</td>
</tr>
<tr>
<td>Max.</td>
<td>3.88</td>
<td>12.01</td>
<td>129.88</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.94</td>
<td>1.98</td>
<td>21.92</td>
</tr>
</tbody>
</table>
3. Soft sensor structure

The aim of implementing a soft sensor for tomato crops in greenhouses is to offer an alternative method to destructive tests in order to certainly determine biomass values for crop growth status. In combination with IoT systems, a soft sensor is intended as a powerful tool for continuous crop monitoring.

The proposed soft sensor (Fig. 8) is based on the tomato growth model previously described. Therefore, LAI can be estimated indirectly from other measurements provided by real sensors installed inside the greenhouse. Also, electrical conductivity should be considered as an input for the soft sensor, although in this work has been omitted because it was automatically controlled through fertigation at a fixed value.

3.3 Short term prediction mechanism

A prediction mechanism has been proposed to offer a future estimation of tomato crop evolution by combining the greenhouse climate model and the tomato growth model as a soft sensor. The complete structure of this mechanism is presented in Fig. 9. The mechanism performs data assimilation from three different sources: (i) AEMET weather forecasts every 6 hours, (ii) measured data from greenhouse real sensors, and (iii) real biomass samples when available during a season.

This prediction mechanism is recursive and is based on a sequence that allows to execute two differentiated computational loops. The main loop is dedicated to predicting crop growth with a 48 hours horizon for each AEMET weather forecast. The secondary loop is exclusively executed to estimate crop growth by utilizing registered past data from inside greenhouse sensors, accumulated during the last 6 hours (AEMET forecast intervals). This secondary loop is intended as a correction action for a better Leaf Area Index estimation provided by the designed soft sensor.

Table 5. Error analyses for soil surface temp. sub-model

<table>
<thead>
<tr>
<th></th>
<th>Calibration (°C)</th>
<th>Validation (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variation interval</td>
<td>[21.61, 46.90]</td>
<td>[22.38, 44.26]</td>
</tr>
<tr>
<td>MAE</td>
<td>1.58</td>
<td>2.17</td>
</tr>
<tr>
<td>Max.</td>
<td>4.09</td>
<td>4.47</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.86</td>
<td>0.86</td>
</tr>
</tbody>
</table>
Utilizing real values as inputs for tomato growth soft sensor ensures to eliminate the uncertainty partially originated by the greenhouse climate model and AEMET weather forecasts. Corrected values for tomato growth state variables are fed back to initialize the main loop before its next execution. Greenhouse climate model state variables are also initialized with the last available measurements from real sensors.

Fig. 9. Leaf Area Index soft sensor integrated in the short-term prediction mechanism.

Inside the main loop, a temperature and humidity control algorithm has been implemented to simulate the actual control system of the experimental greenhouse. It consists of a series of heuristics rules (a kind of gain-scheduling control approach) to calculate the opening percentage for the natural ventilation in order to achieve proper climate conditions for the crop inside the greenhouse. It was necessary to include this control block to predict future behaviour for actuators since they affect to greenhouse climate conditions.

As a preliminary stage to utilize the complete prediction mechanism, AEMET weather forecasts should be evaluated in conjunction with the greenhouse climate model. It is important to highlight that AEMET provided weather forecasts for coordinates corresponding to the University of Almería, which is located 25 km far from the experimental greenhouse. Thus, it is expected to observe some errors between forecasts and real external measurements.

Fig. 10 shows a comparison of an AEMET weather forecast (represented in local time: 02:00 GMT + 48 h horizon) and real data measured outside the greenhouse. External global radiation is predicted with great accuracy on sunny days. External air temperature and relative humidity have quite similarity in tendency, although relative humidity presents a substantial error in the first hours. The biggest error is observed for wind velocity. This forecast was then utilised as inputs to simulate climate model, obtaining results offered in Fig. 11. Climate model state variables are plotted with a surrounding confidence area corresponding to MAE values that each state variable presented in the model validation stage. In general, an adequate behaviour is observed for all state variables, despite an overestimation for inside air temperature in the middle of each day, or an underestimation at night periods for soil surface temperature. The inside air relative humidity also presents an overestimation in the first hours of August 23, while the rest of the prediction clearly fits to real tendency. Mentioned differences may be justified considering existing mismatches between external weather forecast and real climate conditions.

Once the quality of climate predictions was evaluated, the complete short-term prediction mechanism was tested. To that end, the tomato growth soft sensor was previously executed with registered data during 15 days until September 8, when LAI was estimated in 0.5 (m² leaf/m² ground). Unfortunately, real LAI samples were not available to confirm the accuracy of the soft sensor. However, the estimated value was close to historical registers for other seasons beginning in similar dates. Hence, the estimated value for LAI was utilised to initialize the short-term prediction mechanism at September 8.
In order to test the recursive action, two AEMET weather forecasts were utilised for the same day, as seen in Fig. 12. Results of inside climate state variables are presented in Fig. 13, in which outstanding differences between predictions and real data are appreciable. These discrepancies directly affect to LAI predictions, as shown in Fig. 14. In this graphic, the blue line could be considered the “true” measurement for LAI, since it was generated by the soft sensor. The range of variation for LAI is small in a period of 48 hours as crop growth is a slow dynamic process. Even so, it can be noted that for the prediction beginning at 02:00 (local time), the predicted LAI progressively drifts apart from real tendency. Thanks to correction action at 14:00 (local time), LAI is reinitialized to the true value and the second prediction starts. A few hours later (in September 9), error between second prediction and soft sensor tendency is increasing. This justifies the necessity of recursively correct each prediction with the execution of the tomato growth soft sensor.

![Fig. 12. AEMET weather forecasts as inputs for the short-term prediction mechanism.](image1)

![Fig. 13. Predictions for inside greenhouse climate as a result for the short-term prediction mechanism.](image2)

![Fig. 14. LAI predictions compared against soft sensor tendency.](image3)

In this work, a soft sensor to estimate tomato crop growth has been proposed. The main function of the soft sensor is to obtain an estimation for LAI suppressing the need to manually perform several destructive tests.

The developed methods to predict tomato crop growth demonstrate a reasonable performance under different dynamics. These methods have a clear application to real scenarios currently presented by greenhouses in the province of Almería (Spain). The proposed tomato growth soft sensor could be directly applied to those greenhouses with inside installed sensors or IoT systems. Other greenhouses with low level of technology or lack of inside sensors should utilise the short-term prediction mechanism as a support tool for decision making, but in this case, correction for the predicted LAI tendency still must be carried out by periodically registering real biomass samples.

Future works may be aimed at introducing new dynamic terms to improve the utilised models and the developed soft sensor. For instance, simulation for CO₂ concentration in the air could be incorporated to the greenhouse climate model. Moreover, the effect of vapor pressure deficit and water deficit on the crop should be consider for the tomato growth model. Additionally, the proposed prediction mechanism could be complemented with artificial vision techniques based on regularly taking pictures of the crop.

ACKNOWLEDGEMENTS

This work has been funded by the following projects: DPI2017-85007-R (Spanish Ministry of Science, Innovation and Universities and ERDF funds), and the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 731884, IoF2020. Author F. García-Mañas is supported by an FPU grant of the Spanish Ministry of Science, Innovation and Universities. Authors also appreciate the collaboration of Cajamar Foundation.

This manuscript contains results obtained from the information provided by the State Meteorological Agency (AEMET), Spanish Ministry of Agriculture, Food and Environment.
REFERENCES


