Chapter 2
The Greenhouse Dynamical System

2.1 Climate Dynamic Models

As pointed out in [324], when a complex system is modeled, one of the questions that arises is to discern whether models based on first principles or empirical models based on experimental data are to be used. The former generally provide detailed information of the process than empirical models, but they are usually more complex requiring longer times and deep knowledge in the design phase. Although models based on first principles can be used within model-based control structures, they are generally used for simulation purposes, while empirical ones are used for control tasks. These two approaches (and combinations) can be found within the framework of greenhouse climate variables modeling. This section presents the development of climate dynamic models based on first principles and on input–output data.

2.1.1 First Principles-Based Models

2.1.1.1 General Considerations

The dynamic behavior of the microclimate inside a greenhouse is a combination of physical processes involving energy transfer (radiation and heat) and mass balance (water vapor and CO_2 fluxes). These processes depend on the outside environmental conditions, structure of the greenhouse, type and state of the crop, and on the effect of the control actuators (typically ventilation and heating to modify inside temperature and humidity conditions, shading and artificial lighting to change internal radiation, CO_2 enrichment to influence photosynthesis and fogging/cooling for humidity enrichment).

The development of models of a dynamic system is a complex process that depends on the characteristics of the dynamics of the process object of study. This section deals with models based on physical principles as this is not a completely solved

F. Rodríguez et al., *Modeling and Control of Greenhouse Crop Growth*,
Advances in Industrial Control, DOI 10.1007/978-3-319-11134-6_2

problem. These models have been developed in different parts of the world since the 1960s of the last century, applied to several greenhouse structures (many of them of small size and used for research purposes), with different climatic actuators, cover material, and crops. Among them, it is interesting to emphasize several works related to that presented in this chapter, performed by the following authors:

- North and Central Europe: Bot [57], Udink ten Cate [461], Halleaux [167], Young et al. [483], van Henten [177], Tchamitchan et al. [447, 448], Tap et al. [441–443], Speetjens et al. [416]. It is interesting to highlight the work of Vanthoor et al. [468], in which a general methodology for any latitude is developed, using a similar approach as that proposed in this book.
- Mediterranean area: Kindelan [217], Cormary and Nicolas [99], Chaabane [83], Manera et al. [264], Boisson [55], Ioslovich el al. [202], Boulard et al. [62], Zhang et al. [489], Senent et al. [384], Wang and Boulard [476], and Tavares et al. [446].
- Central and North America: Takakura el al. [436, 437], Ahmadi et al. [3–5], Halleaux [167], Trigui et al. [459, 460], Leal-Iga [244], Bot et al. [198].
- South of Asia: Sharme et al. [392].

Although all these models are based on the same physical principles, they show differences in the approaches used when adapted to the particular conditions in each area. All these works describe the basic equations of the mathematical models and include some results, but they do not describe the complete methodology used for the implementation, calibration, and validation of the models. Other approaches can be found in [140, 306].

To model the climate that is generated inside a greenhouse based on physical, physiological, biological, and chemical principles, mass and energy balances have to be applied to all its constitutive elements. The main subsystems are [305]:

- *Cover*: It is a solid and homogeneous medium which partially transmits solar and thermal radiation. Its main objective is to isolate the internal atmosphere of the external weather conditions, making a bridge between the two environments.
- *Crop*: It is a living organism that is an open thermodynamic system that extracts energy from the surrounding environment to create and maintain its own essential management.
- *Air*: It is a gaseous medium joining the different solid elements in the greenhouse.
- *Soil*: It is a porous medium in which can be distinguished a solid phase (soil and organic matter), a liquid phase (water), and a gaseous phase (vapor and air). It is responsible for the greenhouse thermal inertia, absorbing energy during the day and emitting it overnight. Actually, it is divided into:
  - Surface, that is, the interface with the rest of the greenhouse.
  - Lower layers, that separate layers of ground that have different thermal characteristics related by conduction processes.

Therefore, the variables that describe the greenhouse climate are: Air temperature, water content in the air, CO_2 concentration in the air, temperatures of the outer and

inner surface of the cover, crop temperature, soil surface temperature, and temperature of each of the layers in which the soil is divided. Among these elements, the various energy and mass transport processes occur (conduction, convection, radiation, condensation, evaporation, and transpiration). Moreover, these processes are affected by other climatic variables as air speed inside the greenhouse and apparent temperature of the sky, which is defined as the temperature of a black hemisphere exchanging thermal radiation with the different elements of the greenhouse according to the Stefan–Boltzmann law, in the same amount as the actual exchange that occurs between the greenhouse and the atmosphere [58].

Other devices to consider when modeling greenhouse climate are installed actuators (those used to modify climatic variables) that constitute the inputs to the system and that can be artificially manipulated. As discussed above, there is a wide variety of climatic actuators, although the most common in warm climates are natural ventilation, heating systems, shading and thermal screens, humidifiers, and CO_2 enrichment systems.

In a greenhouse, the Principle of Continuity between its elements applies [278], so that the heat and mass transfer processes in each can be studied using mass and energy balances.

The energy balance in a given volume (vol) is described by the following differential equation:

$$\frac{\mathrm{d}Q_{\mathrm{tot,vol}}}{\mathrm{d}t} = c_{\mathrm{h,vol}} \frac{\mathrm{d}X_{\mathrm{T,vol}}}{\mathrm{d}t} = Q_{\mathrm{in,vol}} - Q_{\mathrm{out,vol}} + Q_{\mathrm{gen,vol}} \tag{2.1}$$

where $Q_{\mathrm{tot,vol}}$ (J) is the total amount of energy accumulated in the volume, $Q_{\mathrm{in,vol}}$ and $Q_{\mathrm{out,vol}}$ (J s^{-1}) are the energy per time unit entering and leaving the volume, respectively, and $Q_{\mathrm{gen,vol}}$ (J s^{-1}) is the energy generated inside the volume. The left-hand term represents the change in energy per time unit t in the considered volume, which is directly related to temperature, $X_{\mathrm{T,vol}}$ (K), through the heat capacity $c_{\mathrm{h,vol}}$ (J K^{-1}).

The same considerations can be done with the mass balances in a volume (related with concentration, $X_{\mathrm{c,vol}}$ (kg m^{-3}) and volume c_{vol} (m^3)), in such a way that the variation with time of the mass within a determined volume, $M_{\mathrm{tot,vol}}$ (kg), is equal to the difference between the input, $M_{\mathrm{in,vol}}$ and output, $M_{\mathrm{out,vol}}$, flows (kg s^{-1}), plus the mass generated per time unit inside the volume, $M_{\mathrm{gen,vol}}$ (kg s^{-1}), following the next balance:

$$\frac{\mathrm{d}M_{\mathrm{tot,vol}}}{\mathrm{d}t} = c_{\mathrm{vol}} \frac{\mathrm{d}X_{\mathrm{c,vol}}}{\mathrm{d}t} = M_{\mathrm{in,vol}} - M_{\mathrm{out,vol}} + M_{\mathrm{gen,vol}} \tag{2.2}$$

Therefore, greenhouse climate is defined by a system of ordinary differential equations (ODEs) describing the mass and energy balances:

- Energy balances in: Outer and inner surfaces of the cover, inside air, crop, soil surface, and soil layers (typically from 2 to 5).
- Mass balances: Of water vapor and CO_2 concentration in the greenhouse air.

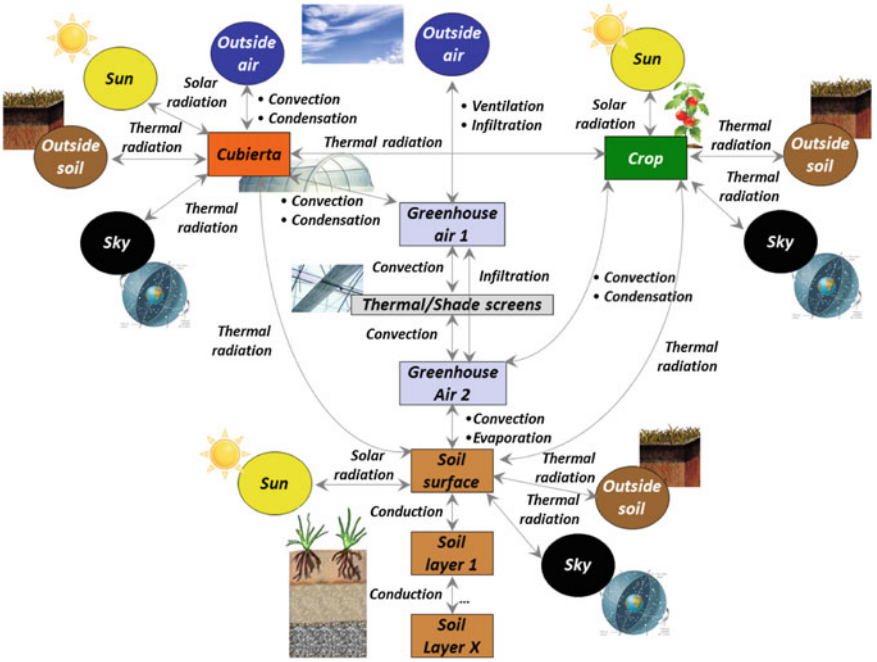

Fig. 2.1 Relationship between greenhouse elements

The number of equations to be solved depends on the known or measured variables, that is, on the boundary conditions. All authors agree to adopt as boundary conditions all the greenhouse climate disturbances, i.e., outdoor climate, soil temperature at a given depth, and wind speed inside. As the interest is in modeling inside air variables (temperature, humidity, and CO_2 concentration), if the variables describing the other elements are measured, they can be used as boundary conditions, thus reducing the complexity of the modeling problem as the number of ODEs is reduced. However, due to technical or economic reasons, sometimes several of these variables are not measured, being necessary to estimate them.

Therefore, to model climate variables in the volume of air that is in direct contact with the crop, the system is divided into the following elements: Cover, crop, soil surface, soil layers, and volume of air between the cover and the ground surface. If a shading screen is installed, the volume of air between the cover and the ground surface is divided into the corresponding two air volumes. Moreover, the surrounding conditions of the system are defined by four elements: Sun, sky dome, outside air, and ground outside the greenhouse.

Among these elements, energy transport phenomena are produced by heat transfer (conduction, convection, absorption, reflection and transmission of solar radiation, emission, absorption, reflection, and transmission of thermal radiation), mass transfer

(condensation of water vapor, evaporation of water vapor, crop transpiration) and the effects of. actuation systems

To summarize, the relationship among the most used actuation systems, the elements of the greenhouse and the outer systems are:

- *Natural ventilation*: It affects the thermal, vapor, and CO_2 balances in indoor air as it mixes with the outside.
- *Shade screen*: It reduces the amount of radiation that reaches the crop and the soil surface. A convection process between the air and the surfaces of the shade occurs. It also may produce condensation phenomena on its surfaces. Finally, as it is composed of porous material, an infiltration phenomenon occurs between the two volumes of air it defines.
- *Thermal screen*: It is a less porous element than the shade screen and also reduces the loss of thermal radiation from the ground and crop.
- *Heating*: If hot water pipes are used (see Sect. 3.1.2.2), convection processes with the surrounding air and thermal radiation exchange processes with soil, crop, cover, screen, outside ground, and sky dome occur.
- *Humidifiers*: They increase the concentration of water vapor in the air, they cause a reduction in the temperature therein.
- CO_2 *enrichment systems*. They increase the concentration of CO_2 in the air.

The dynamics of the climatic variables in a greenhouse are complex due to the following facts [361, 363]:

- Presence of different timescales, from minutes to months.
- Presence of nonlinearities, both static and dynamic.
- Time-varying parameters.
- The system is subjected to strong disturbances (measurable and nonmeasurable ones).
- High degree of correlation among variables.
- Combination of continuous and discrete variables.
- Presence of unmodeled dynamics.
- Changing dynamics depending on the greenhouse characteristics and geographical area.

It is thus a complex system and, although the physical processes taking place in a greenhouse are known, a number of assumptions have to be made to simplify the problem. The hypotheses accepted by most authors are [305]:

- *Cover*: Its material is homogeneous, with constant thermodynamic and optical properties and negligible heat capacity. A descriptive temperature is considered on each side.
- *Crop*: It is a subsystem with uniform density of vegetation that absorbs and transmits solar and thermal radiation. Its thermal capacity can be considered negligible and uniform temperature throughout its volume is assumed.
- *Air*: It is considered homogeneous in terms of thermodynamic properties except in models that include forced ventilation.

- *Soil*: It is considered as a medium divided into a finite number of horizontal layers which are assumed homogeneous in their thermodynamic properties and chemical composition. Heat flux is generally considered unidirectional, regardless of the study of water movement.

2.1.1.2 First Principles Model Architecture and General Hypotheses

To model the distributed nature of the greenhouse, a partial differential equation (PDE) model should be used to account for both time and spatial evolution of the state variables of the system. Nevertheless, greenhouses are often equipped with few sensors and the actuators affect all the greenhouse volume, so that a typical assumption is to consider a perfect mixing behavior such that the greenhouse dynamics are defined by a system of ODEs given as

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = f(\mathbf{X}, \mathbf{U}, \mathbf{D}_m, \mathbf{V}, \mathbf{C}, t) \text{ with } \mathbf{X}(t_i) = \mathbf{X}_i \tag{2.3}$$

where $\mathbf{X} = \mathbf{X}(t)$ is a n-dimensional vector of state variables, $\mathbf{U} = \mathbf{U}(t)$ is a m-dimensional vector of input variables, $\mathbf{D}_m = \mathbf{D}_m(t)$ is an o-dimensional vector of measurable disturbances, $\mathbf{V} = \mathbf{V}(t)$ is a p-dimensional vector of system variables, $\mathbf{C}$ is a q-dimensional vector of system constants, t is the time, $\mathbf{X}_i$ is the known initial state at the initial time t_i and $f = f(t)$ is a nonlinear function based on mass and heat transfer balances.

The number of equations describing the system and their characteristics depend on the greenhouse elements, the installed control actuators, and the type of cultivation method. The model presented in this section corresponds to a typical greenhouse located in the Mediterranean area with a tomato crop. It has been developed assuming some general hypotheses:

- The greenhouse is divided into four elements (Fig. 2.2): Cover, internal air, soil surface, and one soil layer. The crop is not considered as an element as no measurements of the leaf temperature are usually available (the related sensors are not very accurate) and thus it is considered as a source of disturbance for the inside climate. As some of the physical processes require the crop temperature to be known (i.e., thermal radiation among the solid elements), it has been considered to be equal to the greenhouse air temperature.
- The state variables of the model are: The internal air temperature ($X_{\mathrm{T,a}}$) and humidity (absolute $X_{\mathrm{H_{a,a}}}$ and relative $X_{\mathrm{H_r,a}}$), cover temperature ($X_{\mathrm{T,cv}}$), soil surface temperature ($X_{\mathrm{T,ss}}$), and first soil layer temperature ($X_{\mathrm{T,sl}}$). The PAR radiation onto the canopy (output variable $X_{\mathrm{rp,a}}$) is also modeled with an algebraic equation. The CO_2 concentration is measured.
- The exogenous and disturbance inputs acting on the system are the outside air temperature ($D_{\mathrm{T,e}}$) and absolute humidity ($D_{\mathrm{H_{a,e}}}$), wind speed ($D_{\mathrm{ws,e}}$) and direction ($D_{\mathrm{wd,e}}$), sky temperature ($D_{\mathrm{T,sky}}$), calculated using the Swinbank formula [55],

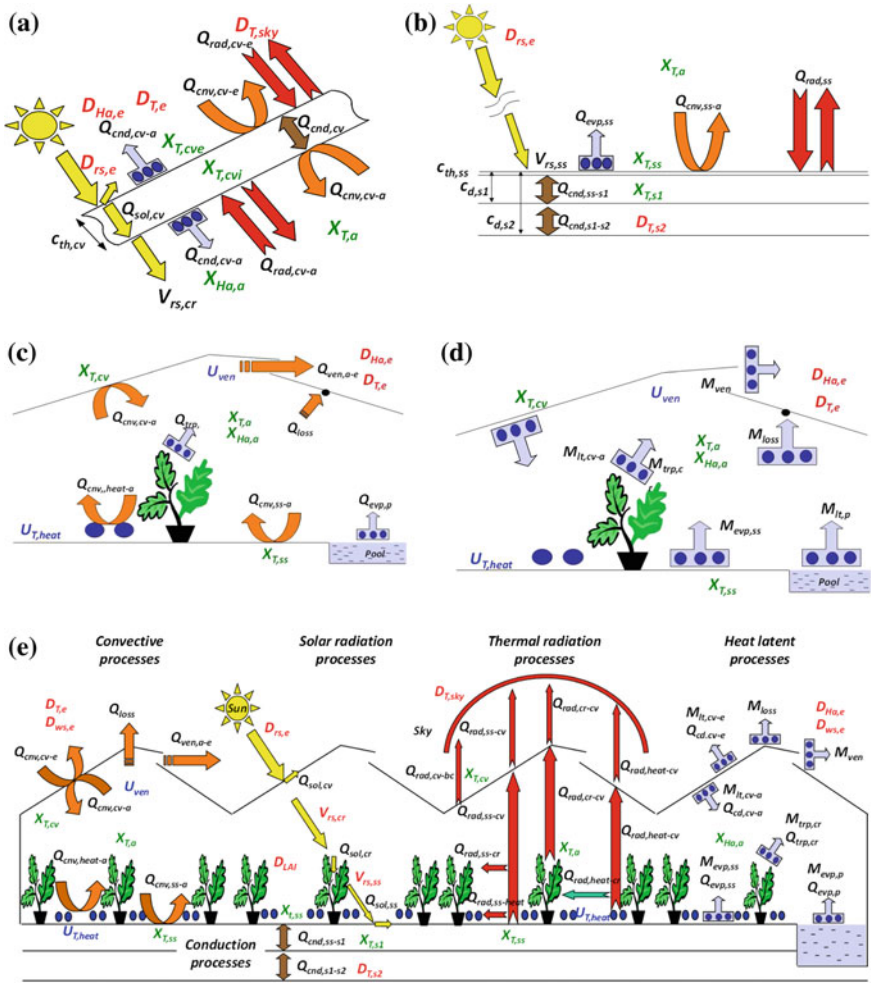

Fig. 2.2 Heat and mass transfer fluxes in a greenhouse. **a** Heat transfer fluxes in a cover. **b** Heat transfer fluxes in the soil layers. **c** Heat transfer fluxes with the internal air. **d** Mass transfer fluxes with the internal air. **e** Complete heat and mass transfer fluxes with the internal air

outside global solar radiation ($D_{\mathrm{rs,e}}$), PAR radiation ($D_{\mathrm{rp,e}}$), greenhouse whitening (D_{wh}) [26], the transpiration rate inside the greenhouse via the leaf area index (LAI, D_{LAI}) and the temperature of the deepest soil layer ($D_{\mathrm{T,s2}}$) which can be calculated as the average of the external air temperature during one year or measured using dedicated sensors [55].

- The control inputs of the system are the position of the natural ventilation (U_{ven}), the position of the shade screen (U_{shd}) and the heating system control signal ($U_{\mathrm{T,heat}}$, that is the temperature of the water of the pipes or the air heater status, depending on the type of heating system used, as commented in Sect. 3.1.2.2).

- The heat fluxes are one-dimensional. The model only considers the vertical dimension.
- The temperature models are based on a heat transfer balance where the following physical processes are included: Solar (sol) and thermal radiation (rad) absorption, heat convection (cnv) and conduction (cnd), crop transpiration (trp), condensation (cd), and evaporation (evp).
- In order to design the humidity model, a mass balance is used based on artificial water influxes, exchange with the outside, crop, condensation, and evaporation.
- The models of short and long wave radiation do not consider reflection, and the air is inert to these processes.
- The physical characteristics of the different elements (cover material, soil components, air, etc.), such as density or specific heat are considered constant in the temperature range the greenhouse evolves.
- The thickness of the cover is in microns, so the conductive heat flow is quantitatively negligible compared to other heat flows that appear in the cover temperature models. For this reason, it is accepted that the temperatures of both cover surfaces are similar.

In what follows, the models representing the heat transfer and mass balances in the four elements constituting the greenhouse are developed. The units of the different variables are indicated in the acronyms section.

2.1.1.3 Model of the PAR Radiation

The PAR radiation onto the canopy is modeled using an algebraic equation, because it is similar to the PAR radiation outside the greenhouse dimmed by the different physical elements that absorb the radiation (mainly cover material, cover whitening and shade screen). So it is modeled using Eq. (2.4).

$$X_{\mathrm{rp,a}} = V_{\mathrm{tsw,g}} D_{\mathrm{rp,e}} \tag{2.4}$$

where $V_{\mathrm{tsw,g}}$ is the greenhouse short wave radiation transmission coefficient, described by:

$$V_{\mathrm{tsw,g}} = \begin{cases} c_{\mathrm{tsw,cv}} & \text{no shade, no whitening} \\ c_{\mathrm{tsw,cv}} c_{\mathrm{tsw,wh}} & \text{no shade, whitening} \\ c_{\mathrm{tsw,cv}} c_{\mathrm{tsw,shd}} & \text{shade, no whitening} \\ c_{\mathrm{tsw,cv}} c_{\mathrm{tsw,wh}} c_{\mathrm{tsw,shd}} & \text{shade, whitening} \end{cases} \tag{2.5}$$

where $c_{\mathrm{tsw,cv}}$ is the cover solar transmission coefficient, $c_{\mathrm{tsw,shd}}$ is the shade screen solar transmission coefficient and $c_{\mathrm{tsw,wh}}$ is the whitening solar transmission. This last parameter is difficult to determine because it depends on the whitening concentration between 4 kg whitening/4 l water ($c_{\mathrm{tsw,wh}} = 0.1$) and 0.7 kg whitening/4 l water ($c_{\mathrm{tsw,wh}} = 0.65$) [278]. It is thus necessary to take measurements of global and PAR

radiation inside and outside the greenhouse to determine this coefficient. Another option is to search the value of this parameter in the modeling calibration phase. Note that Eq. (2.5) introduces a switch in the simulation process. As will be discussed later, the simulation packages used (both block-oriented ones as Simulink [268] and object-oriented ones as Modelica [117]) can cope with such behavior. The same happens with Eqs. (2.14), (2.19) and (2.22).

2.1.1.4 Heat Transfer Through the Cover

As Fig. 2.2a shows, the cover has two sides with different temperatures. Due to the fact that the cover is made using a single material (plastic film) and that its thickness is a few microns, the conduction heat flux, $Q_{,\mathrm{cv}}$, is quantitatively not significant compared to the other fluxes appearing in the balance given by Eq. (2.6) [138]. So, the temperatures of the two sides are assumed to be similar and only one cover temperature is modeled ($X_{\mathrm{T,cv}}$) using the heat transfer balance given by Eq. (2.6).

$$c_{\mathrm{sph,cv}} c_{\mathrm{den,cv}} \frac{\mathrm{c_{vol,cv}}}{\mathrm{c_{area,ss}}} \frac{\mathrm{d}X_{\mathrm{T,cv}}}{\mathrm{d}t} = Q_{\mathrm{sol,cv}} - Q_{\mathrm{cnv,cv-a}} - Q_{\mathrm{cnv,cv-e}} - Q_{\mathrm{cd,cv}} + Q_{\mathrm{rad,cv}} \tag{2.6}$$

where $Q_{\mathrm{sol,cv}}$ is the solar radiation absorbed by the cover, $Q_{\mathrm{cnv,cv-a}}$ is the convective hear transfer with the internal air, $Q_{\mathrm{cnv,cv-e}}$ is the convective heat transfer with the outside air, $Q_{\mathrm{cd,cv}}$ is the latent heat produced by condensation on both sides of the cover, $Q_{\mathrm{rad,cv}}$ is the thermal radiation absorbed by the cover from the inside and outside of the greenhouse, $c_{\mathrm{sph,cv}}$ is the specific heat of the cover material, $c_{\mathrm{den,cv}}$ is the cover material density, $c_{\mathrm{vol,cv}}$ is the cover volume and $c_{\mathrm{area,ss}}$ is the greenhouse soil surface.

The solar radiation absorbed by the cover is determined by the short wave radiation cover material absorptivity, $c_{\mathrm{asw,cv}}$, using the following equation:

$$Q_{\mathrm{sol,cv}} = c_{\mathrm{asw,cv}} D_{,\mathrm{e}} \tag{2.7}$$

The convective heat transfer from inside air to cover is calculated based on the difference between the cover temperature, $X_{\mathrm{T,cv}}$, and the greenhouse air temperature, $X_{\mathrm{T,a}}$, using the typical model of this type of heat transfer:

$$Q_{\mathrm{cnv,cv-a}} = V_{\mathrm{cnv,cv-a}} \frac{\mathrm{c_{area,cv}}}{\mathrm{c_{area,ss}}} \left(X_{\mathrm{T,cv}} - X_{\mathrm{T,a}}\right) \tag{2.8}$$

where $c_{\mathrm{area,cv}}$ is the cover surface, $V_{\mathrm{cnv,cv-a}}$ is the cover inside convective heat transfer coefficient based on the difference between the cover temperature and the internal air temperature, and the mean greenhouse air speed, $V_{\mathrm{ws,a}}$:

$$V_{\mathrm{cnv,cv-a}} = c_{\mathrm{cnv,cv-a1}} \left|X_{\mathrm{T,cv}} - X_{\mathrm{T,a}}\right|^{c_{\mathrm{cnv,cv-a2}}} + c_{\mathrm{cnv,cv-a3}} \left(V_{\mathrm{ws,a}}\right)^{c_{\mathrm{cnv,cv-a4}}} \tag{2.9}$$

where $c_{\mathrm{cnv,cv-ax}}$ are empirical parameters that have to be estimated. This analysis uses the Nusselt, Prandtl, Grashof, and Reynolds numbers related to the climate variables involved in this process. There are tables with general cases, facilitating the calculations. The parameters $c_{\mathrm{cnv,cv-a1}}$ and $c_{\mathrm{cnv,cv-a2}}$ are different depending on the convection type (laminar or turbulent). In order to simplify the model, the approach proposed by Chalabi and Bailey [85] is used: If the internal air temperature is higher than the cover temperature, the heat transfer is turbulent; otherwise it is laminar. On the other hand, the parameters $c_{\mathrm{cnv,cv-a1}}$ and $c_{\mathrm{cnv,cv-a3}}$ vary with the position of the shade screen. When the screen is extended, the air is divided into two volumes, so it is necessary to include three new balance equations (air between the cover and the screen, upper and lower surfaces of the screen). Measurements of these surface temperatures are not usually available, so the effect of the shade screen on the convective coefficient is modeled by decreasing the value of this parameter. As will be seen in the next sections, good results are obtained under this simplification.

The measurement of the greenhouse air speed is a difficult task, because during long time intervals of the greenhouse operation the values are very low ($< 1\,\mathrm{ms}^{-1}$). So it is necessary to use special anemometers (like ultrasound or thermal effect based ones). As the installation of such sensors is not usual in Mediterranean greenhouses, it can be estimated using the studies in [477], which provide the following expression:

$$V_{\mathrm{ws,a}} = \frac{V_{\mathrm{ven,flux}}}{c_{\mathrm{ven,areap}}} \tag{2.10}$$

where $c_{\mathrm{ven,areap}}$ is the greenhouse section area perpendicular to the ventilation flux and $V_{\mathrm{ven,flux}}$ is the volumetric flow rate (also known as ventilation rate). There are different theories to calculate this last variable. Models "M1" and "M4" proposed by Boulard and Baille [61] have been used because the type of greenhouse structures studied are similar to those treated in this book, equipped with long continuous roofs. Moreover, the five models proposed by Boulard and Baille [61] were tested and "M1" and "M4" fix better to the data. These models are based on the thermal buoyancy (depending on the temperature difference between inside and ourside air ($X_{\mathrm{T,a}} - D_{\mathrm{T,e}}$)) and wind forces (function of the outside wind speed $D_{\mathrm{ws,e}}$), and are described by Eqs. (2.11) and (2.12),

$$V_{\mathrm{ven,flux}_{M1}} = \frac{c_{\mathrm{ven,n}} c_{\mathrm{ven,l}} c_{\mathrm{ven,d}} D_{\mathrm{T,e}}}{3 c_{\mathrm{gv}} (X_{\mathrm{T,a}} - D_{\mathrm{T,e}})} \left[\left(V_{\mathrm{ven,hef}} c_{\mathrm{gv}} \frac{X_{\mathrm{T,a}} - D_{\mathrm{T,e}}}{D_{\mathrm{T,e}}} + c_{\mathrm{ven,wd}} D_{\mathrm{ws,e}}^2 \right)^{3/2} - (c_{\mathrm{ven,wd}} D_{\mathrm{ws,e}}^2)^{3/2} \right] + V_{\mathrm{loss}} \tag{2.11}$$

$$V_{\mathrm{ven,flux}_{M4}} = \frac{c_{\mathrm{ven,n}} c_{\mathrm{ven,l}} c_{\mathrm{ven,d}} V_{\mathrm{ven,hef}}}{2} \left[\left(c_{\mathrm{gv}} \frac{V_{\mathrm{ven,hef}}}{2} \frac{X_{\mathrm{T,a}} - D_{\mathrm{T,e}}}{D_{\mathrm{T,e}}} \right)^{0.5} + (c_{\mathrm{ven,wd}}^{0.5} D_{\mathrm{ws,e}}) \right] + V_{\mathrm{loss}} \tag{2.12}$$

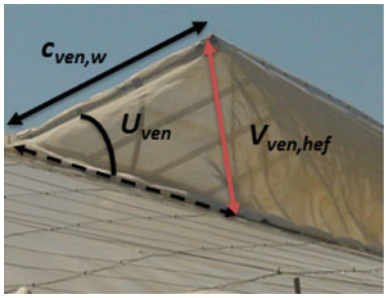

Fig. 2.3 Relationship between vents aperture and effective height of ventilation

where $c_{\text{ven,n}}$ is the number of vents, $c_{\text{ven,l}}$ is the length of the vents, $c_{\text{ven,d}}$ is the discharge coefficient, c_{gv} is the gravity constant, $c_{\text{ven,wd}}$ is the wind effect coefficient, and $V_{\text{ven,hef}}$ is the cord joining the two extremities of the vent based on the position of the vent [rad, °], U_{ven}, using the following equation (see Fig. 2.3):

$$V_{\text{ven,hef}} = 2c_{\text{ven,w}} \sin\left(U_{\text{ven}}/2\right) \tag{2.13}$$

where $c_{\text{ven,w}}$ is the width of vent.

V_{loss} is the leakage when the vent is closed, based on the wind speed, which can be approximated by:

$$V_{\text{loss}} = \begin{cases} c_{\text{loss,lw}} & D_{\text{ws,e}} < c_{\text{ws,lim}} \\ c_{\text{loss,hw}} & D_{\text{ws,e}} \geq c_{\text{ws,lim}} \end{cases} \tag{2.14}$$

$c_{\text{loss,lw}}$ being the leakage with low wind speed, $c_{\text{loss,hw}}$ is the leakage with high wind speed and $c_{\text{ws,lim}}$ is the wind speed considered as the limit between high and low wind. In [61], the authors proved empirically that the discharge and wind effect coefficients are not really constant and their values depend on some variables as the wind speed, but in this work they are considered to be constant due to the difficulty involved in estimating these relations. After calibration of the model, the values obtained for these parameters were lower than those provided by the references due to the effect of insect-proof screens located on the vents [278]. A study was also performed to analyze the effect of wind direction modifying the structure of the model. The wind speed was modulated based on direction and orientation of vents and it was observed that the wind effect was low, dependent on wind direction. This result agrees with the conclusion drawn by Boulard and Baille [61]. In the case that the greenhouse has lateral and roof ventilation, the following expression can be used to estimate ventilation rate [219]:

$$V_{\text{ven,flux}} = c_{\text{ven,d}}\left[\left(\frac{V_{\text{ven,area-lat}}V_{\text{ven,area-roof}}}{\sqrt{V_{\text{ven,area-lat}}^2+V_{\text{ven,area-roof}}^2}}\right)^2\left(2c_g c_{\text{ven,h}}\frac{X_{\text{T,a}}-D_{\text{T,e}}}{D_{\text{T,e}}}\right)+\left(\frac{V_{\text{ven,area-lat}}V_{\text{ven,area-roof}}}{2}\right)^2 c_{\text{ven,wd}}D_{\text{ws,e}}^2\right]^{0.5}+V_{\text{loss}} \tag{2.15}$$

where $c_{\text{ven,h}}$ is the vertical distance between the midpoints of the lateral and roof vents, $V_{\text{ven,area-lat}}$ and $V_{\text{ven,area-roof}}$ are the areas of the roof and sidewall ventilation openings, given by the following equations based on U_{vent} expressed in %:

$$V_{\text{ven,area-lat}} = c_{\text{ven,l-lat}}c_{\text{ven,w-lat}}(U_{\text{ven}}/100) \tag{2.16}$$

$$V_{\text{ven,area-roof}} = 2c_{\text{ven,l-roof}}c_{\text{ven,w-lat}}\sin\left(\frac{U_{\text{ven}}}{100}\frac{U_{\text{ven,max}}}{2}\right) \tag{2.17}$$

where $c_{\text{ven,l-[lat,roof]}}$ and $c_{\text{ven,w-[lat-roof]}}$ are, respectively, the length and width of lateral or roof vents.

The convective heat transfer from outside air to cover is calculated in a similar way as the inside convective term using the formula:

$$Q_{\text{cnv,cv-e}} = V_{\text{cnv,cv-e}}\frac{c_{\text{area,cv}}}{c_{\text{area,ss}}}\left(X_{\text{T,cv}} - D_{\text{T,e}}\right) \tag{2.18}$$

where $V_{\text{cnv,cv-e}}$ is the cover outside convective heat transfer coefficient based on the difference between the cover temperature and the external air temperature, $D_{\text{T,e}}$, and on the outside wind speed. In this case, the wind effect is predominant, so the temperature effect is neglected in the calculation of $V_{\text{cnv,cv-e}}$. Some authors [57] propose a linear relationship with the wind speed and others [21] propose an exponential one. Both approaches are tested in this work and the data fixed better using a mixed formula including a linear equation for low wind velocity and an exponential equation for high wind speed conditions. This formula is used by other authors as indicated by Boisson [55]:

$$V_{\text{cnv,cv-e}} = \begin{cases} c_{\text{cnv,cv-e1}}D_{\text{ws,e}}^{c_{\text{cnv,cv-e2}}} & D_{\text{ws,e}} > c_{\text{ws,lim}} \\ c_{\text{cnv,cv-e3}}D_{\text{ws,e}} + c_{\text{cnv,cv-e4}} & D_{\text{ws,e}} \leqslant c_{\text{ws,lim}} \end{cases} \tag{2.19}$$

where $c_{\text{cnv,cv-ex}}$ are empirical parameters that have to be estimated.

The most important latent convective fluxes on the cover are produced by condensation on the inside surface. For this reason, some references [277, 447] do not consider the effect of the condensation on the outside surface. Indeed, some authors, van Henten and Tap et al. [177, 443], neglect the effect of condensation on both cover surfaces compared with the other heat processes.

Condensation takes place when water vapor concentration of the internal air, $X_{\text{H}_{\text{a,a}}}$, is greater than water concentration of the cover at saturation, $V_{\text{hsat,cv}}$, calculated based on the cover temperature. This flux can be written as:

$$Q_{\mathrm{cd,cv}} = V_{\mathrm{lt,vap}} M_{\mathrm{cd,cv}} \tag{2.20}$$

where $V_{\mathrm{lt,vap}}$ is the latent heat of evaporation of water calculated at internal air temperature (in °C) using Eq. (2.21).

$$V_{\mathrm{lt,vap}} = 4185.5(597 - 0.56 X_{\mathrm{T,a}}) \tag{2.21}$$

$M_{\mathrm{cd,cv}}$ is the mass condensation flux from the cover calculated based on a convective term:

$$M_{\mathrm{cd,cv}} = \begin{cases} 0 & X_{\mathrm{H_{a,a}}} < V_{\mathrm{hsat,cv}} \\ c_{\mathrm{den,a}} \frac{V_{\mathrm{cnv,cv-a}}}{c_{\mathrm{sph,a}}} \frac{c_{\mathrm{area,cv}}}{c_{\mathrm{area,ss}}} \left(V_{\mathrm{hsat,cv}} - X_{\mathrm{H_{a,a}}} \right) & X_{\mathrm{H_{a,a}}} \geq V_{\mathrm{hsat,cv}} \end{cases} \tag{2.22}$$

where $c_{\mathrm{sph,a}}$ is the specific heat of air and $c_{\mathrm{den,a}}$ is the air density.

The cover thermal radiation flux can be calculated using the Stefan–Boltzmann theory subtracting the thermal radiation emitted by the cover (two surfaces) and the thermal radiation emitted by the other solid elements of the greenhouse: Internal soil surface (ss), pipe heating (heat), crop (cr), and upper hemisphere (sky) that reach the cover surface. Related to the effect of the outside soil surface, some authors consider the temperature similar to the external air temperature [217]. In this proposal, this flux is neglected like other authors (e.g. [447]). The crop is a solid whose surface and volume are variables in time, so the thermal radiation processes between the rest of the solids and the crop are also variable. To model this effect, the long wave crop extinction coefficient, $c_{\mathrm{extlw,cr}}$, and the LAI, D_{LAI}, are used to modulate the crop growth and its effect on thermal processes. The LAI can be measured online or modeled using, for example, Tomgro model developed in [211].

On the other hand, the thermal processes among the soil surface and pipe heating with the rest of the solids are influenced by the crop status because it is located between them, so these processes are modulated by the LAI, so that the heat transfer is smaller when the crop grows. So, this flux can be described by:

$$\begin{aligned} Q_{\mathrm{rad,cv}} = \frac{c_{\mathrm{area,cv}}}{c_{\mathrm{area,ss}}} c_{\mathrm{alw,cv}} c_{\mathrm{sb}} \Big[& \left(c_{\mathrm{vf,ss-cv}} c_{\mathrm{elw,ss}} X_{\mathrm{T,ss}}^4 + c_{\mathrm{vf,heat-cv}} c_{\mathrm{elw,heat}} U_{\mathrm{T,heat}}^4 \right) \\ & \exp\left(-c_{\mathrm{extlw,cr}} D_{\mathrm{LAI}} \right) + c_{\mathrm{vf,sky-cv}} D_{\mathrm{T,sky}}^4 \\ & + c_{\mathrm{vf,cr-cv}} c_{\mathrm{elw,cr}} \left(1 - c_{\mathrm{extlw,cr}} D_{\mathrm{LAI}} \right) X_{\mathrm{T,cr}}^4 - 2 c_{\mathrm{elw,cv}} X_{\mathrm{T,cv}}^4 \Big] \end{aligned} \tag{2.23}$$

where $c_{\mathrm{alw,cv}}$ is the long wave cover absorbance, c_{sb} is the Stefan–Boltzmann constant, $c_{\mathrm{elw,x}}$ are the long wave emissivities of the solid elements, and $c_{\mathrm{vf,x-cv}}$ are the view factors for radiation exchange between the different considered elements x (ss, heat, sky, crs) and the cover. These last parameters can be estimated using input/output data due to the difficulty involved in obtaining their exact values in this type of greenhouse with several surfaces forming the cover and the geometry of the plants elements. In the case that the heating system is based on air heaters, the flux due to this actuator can be ignored.

2.1.1.5 Heat Transfer Fluxes in the Soil Layers

The soil (greenhouse thermal mass) plays an important role in greenhouse climate. During diurnal time, the soil absorbs solar radiation on its surface, heating the deep soil layers. During night, the soil transfers heat to the greenhouse environment from these layers. So, the conductive fluxes are significant because this process is the source of the heat fluxes between them. As shown in Fig. 2.2b, a simple model of the soil is considered, divided into three layers (more layers could be taken into account): Surface, first layer, and a deep layer with a constant temperature. The conduction process is modeled solving the Fourier equation considering one-dimensional heat transfer along the deep axis, in steady state, the different soil layers as flat parallel planes, plus a delay in the process, obtaining acceptable results. This approach is considered because the computational cost decreases when compared with the solution obtained via diffusion equations while the results are similar (see Fig. 2.4).

Soil surface temperature model. Based on energy balance, the temperature of the soil surface (5 cm thickness) is represented by the following equation:

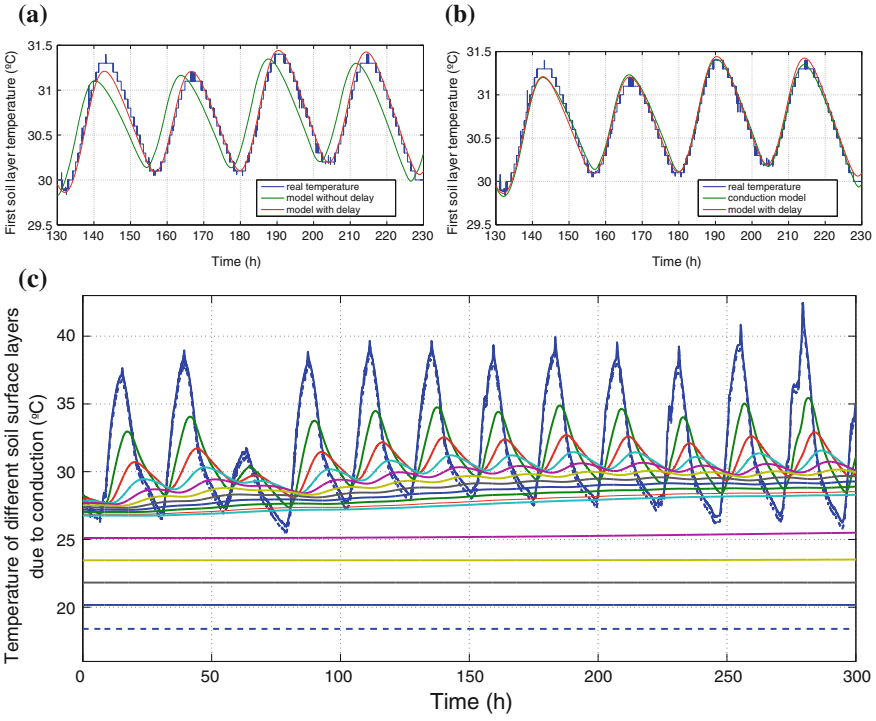

Fig. 2.4 First soil layer temperature model

$$c_{\text{sph,ss}} c_{\text{den,ss}} c_{\text{th,ss}} \frac{\mathrm{d}X_{\text{T,ss}}}{\mathrm{d}t} = Q_{\text{sol,ss}} - Q_{\text{cnv,ss-a}} - Q_{\text{cnd,ss-s1}} - Q_{\text{evp,ss}} + Q_{\text{rad,ss}} \tag{2.24}$$

where $Q_{\text{sol,ss}}$ is the solar radiation absorbed by the soil surface, $Q_{\text{cnv,ss-a}}$ is the convective flux with the internal air, $Q_{\text{cnd,s-s1}}$ is the conductive flux between the soil surface, and the first soil layer located at 30 cm depth, $Q_{\text{evp,ss}}$ is the latent heat produced by evaporation on the soil surface, $Q_{\text{rad,ss}}$ is the thermal radiation absorbed by the soil surface, $c_{\text{sph,ss}}$ is the specific heat of the soil surface material, $c_{\text{den,ss}}$ is the soil surface material density and $c_{\text{th,ss}}$ is the thickness of the soil surface.

The solar radiation absorbed by the soil surface is calculated based on the crop status (defined by LAI), using Eq. (2.25),

$$Q_{\text{sol,ss}} = c_{\text{asw,ss}} V_{\text{rs,cr}} \exp\left(-c_{\text{extsw,cr}} D_{\text{LAI}}\right) \tag{2.25}$$

where $c_{\text{asw,ss}}$ is the solar absorptivity of the soil surface material for short wave radiation, $c_{\text{extsw,cr}}$ is the canopy short wave extinction coefficient and $V_{\text{rs,cr}}$ is the solar radiation that reaches the top of the canopy based on the solar radiation absorption by the physical elements that the radiation crosses:

$$V_{\text{rs,cr}} = V_{\text{tsw,g}} D_{\text{rs,e}} \tag{2.26}$$

$V_{\text{tsw,g}}$ being the greenhouse short wave radiation transmission coefficient defined in Eq. (2.5).

The convective heat transfer from inside air to soil surface is calculated in the same way as cover convective fluxes using the following equation:

$$Q_{\text{cnv,ss-a}} = V_{\text{cnv,ss-a}} \left(X_{\text{T,ss}} - X_{\text{T,a}}\right) \tag{2.27}$$

where $V_{\text{cnv,ss-a}}$ is the inside soil surface convective heat transfer coefficient based on the difference between the soil surface temperature and the internal air temperature, and the mean greenhouse air speed on the soil surface. Using studies of [447], the mean greenhouse air speed proposed is calculated at crop level, so it is modulated based on LAI to obtain an estimation of the greenhouse air speed at soil surface level:

$$V_{\text{cnv,ss-a}} = c_{\text{cnv,ss-a1}} \left| X_{\text{T,ss}} - X_{\text{T,a}} \right|^{c_{\text{cnv,ss-a2}}} + c_{\text{cnv,ss-a3}} \left[V_{\text{ws,a}} \exp\left(c_{\text{cnv,ss-a4}} D_{\text{LAI}}\right) \right]^{c_{\text{cnv,ss-a5}}} \tag{2.28}$$

where $c_{\text{cnv,ss-ax}}$ are empirical parameters that have to be estimated and $V_{\text{ws,a}}$ can be measured or estimated using Eq. (2.10).

The conductive flux between the soil surface and the first soil layer is calculated based on the assumption that the heat flux is one-dimensional (Z axis)

$$Q_{\text{cnd,ss-s1}} = c_{\text{cnd,s1}} \frac{X_{\text{T,ss}} - X_{\text{T,s1}}}{c_{\text{d,s1}} - c_{\text{d,ss}}} \tag{2.29}$$

where $c_{\mathrm{cnd,s1}}$ is the heat conductivity of the first soil layer, $c_{\mathrm{d,ss}}$ is the soil surface deepness, and $c_{\mathrm{d,s1}}$ is the first soil layer thickness.

The latent heat in the soil surface is mainly produced by evaporation, calculated as a convective flux using Eq. (2.30),

$$Q_{\mathrm{evp,ss}} = V_{\mathrm{lt,vap}} M_{\mathrm{evp,ss}} \tag{2.30}$$

where $M_{\mathrm{evp,ss}}$ is the mass evaporation flux from the soil surface, which can be obtained by:

$$M_{\mathrm{evp,ss}} = c_{\mathrm{den,a}} \frac{V_{\mathrm{cnv,ss-a}}}{c_{\mathrm{sph,a}}} \left(V_{\mathrm{hsat,ss}} - X_{\mathrm{H_{a,a}}} \right) \tag{2.31}$$

$V_{\mathrm{hsat,ss}}$ being the water concentration of the soil surface at saturation, calculated based on the soil surface temperature. The diffusion effect to the soil surface of the water content in the internal soil layers is not considered. Some tests were performed to show that this term is negligible when compared with other fluxes due to the fact that the soil surface is mulched [421]. In such cases, evapotranspiration can be considered equal to crop transpiration ($M_{\mathrm{trp,cr}} = V_{\mathrm{ET}}$).

Similar to the cover thermal radiation flux, the Stefan–Boltzmann theory is used to calculate the soil surface thermal radiation flux, considering the effect of the crop growth between the soil surface and the cover and the sky and the effect of the cover long wave transmission, $c_{\mathrm{tlw,cv}}$, in the radiation processes between the soil and the sky. So, the model of this process is as follows:

$$\begin{aligned} Q_{\mathrm{rad,ss}} = c_{\mathrm{alw,ss}} c_{\mathrm{sb}} \Big[& \left(c_{\mathrm{vf,cv-ss}} c_{\mathrm{elw,ss}} X_{\mathrm{T,cv}}^4 + c_{\mathrm{vf,sky-ss}} c_{\mathrm{tlw,cv}} D_{\mathrm{T,sky}}^4 \right) \\ & \exp\left(-c_{\mathrm{extlw,cr}} D_{\mathrm{LAI}}\right) + c_{\mathrm{vf,cr-ss}} c_{\mathrm{elw,cr}} \left(1 - c_{\mathrm{extlw,cr}} D_{\mathrm{LAI}}\right) X_{\mathrm{T,cr}}^4 \\ & - c_{\mathrm{elw,ss}} X_{\mathrm{T,ss}}^4 \Big] \end{aligned} \tag{2.32}$$

where $c_{\mathrm{alw,ss}}$ is the long wave soil surface absorbance, and $c_{\mathrm{vf,x-ss}}$ are the view factors for radiation exchange between the solid elements x and the soil surface and $X_{\mathrm{T,cr}} = X_{\mathrm{T,a}}$ following the hypothesis adopted when developing the model.

Heat transfer fluxes in the first soil layer. As can be seen in Figs. 2.1 and 2.2b, in the first soil layer, only the conductive fluxes are considered and so, the heat balance in this element is represented by Eq. (2.33),

$$c_{\mathrm{sph,s1}} c_{\mathrm{den,s1}} c_{\mathrm{th,s1}} \frac{\mathrm{d}X_{\mathrm{T,s1}}}{\mathrm{d}t} = Q_{\mathrm{cnd,ss-s1}} - Q_{\mathrm{cnd,s1-s2}} \tag{2.33}$$

where $c_{\mathrm{sph,s1}}$ is the specific heat of the first soil layer material, $c_{\mathrm{den,s1}}$ is the first soil layer material density and $c_{\mathrm{th,s1}}$ is the thickness of this layer, $Q_{\mathrm{cnd,ss-s1}}$ is the conductive flux between the soil surface and the first layer of the soil, $Q_{\mathrm{cnd,s1-s2}}$

is the conductive flux between the first soil layer and the deep layer at constant temperature, $D_{T,s2}$, described as

$$Q_{cnd,s1-s2} = c_{cnd,s2} \frac{X_{T,s1} - D_{T,s2}}{c_{d,s2} - c_{d,s1}} \tag{2.34}$$

where $c_{cnd,s2}$ is the heat conductivity of the second soil layer, $c_{d,s2}$ is the second soil layer deep, and $c_{d,s1}$ is the first soil layer deep.

Note that these models are formulated using physical properties of the different materials constituting the soil, as the conductivity coefficient, specific heat, density, or solar absorptivity. As some of these parameters are unknown, they are estimated instead of using approximated values obtained from the literature.

It is interesting to show the behavior of the used simplified model of the soil layer temperature. As indicated, the conduction processes between the different soil layers were modeled considering steady-state regime to solve the Fourier equation. The temperature of the first soil layer was modeled using this approach because there are only conduction processes as energy fluxes. The dynamic response of a soil layer temperature is characterized by a time constant based on the density and specific heat of the material forming the layer and its thickness. Although this approach is commonly used in the literature on greenhouse climate, the model based on the diffusion equation is implemented and calibrated too, to compare the real first soil layer temperature with the temperature estimated by the simplified model (low computational cost) and that estimated by the diffusion model (high computational cost). Figure 2.4a shows that the amplitude of the real temperature is similar to the estimation of the simplified model without delay, although both curves are shifted in the time axis. The delay between the real and the simulated temperature is due to the consideration of the steady-state regime of the heat transfer between the soil layers. On the other hand, if the diffusion equation is solved using Dirichlet conditions in the soil surface and the second soil layer, the estimation of the model is similar in amplitude and delay to the real temperature of the first soil layer (Fig. 2.4b). The considered solution in this work is to use the simplified model including a delay, so that the estimation of the model is similar to the real values, as shown in Fig. 2.4, decreasing the computational cost.

2.1.1.6 Heat Transfer Fluxes with the Internal Air

Based on the processes shown in Fig. 2.2c, the greenhouse air temperature can be modeled using the following balance:

$$c_{sph,a} c_{den,a} \frac{c_{vol,g}}{c_{area,ss}} \frac{dX_{T,a}}{dt} = Q_{cnv,cv-a} + Q_{cnv,ss-a} + Q_{heat-a} - Q_{ven} - Q_{trp,cr} - Q_{evp,p} \tag{2.35}$$

where $Q_{\text{cnv,cv}-\text{a}}$ is the convective flux with the cover described in Eq. (2.8), $Q_{\text{cnv,ss}-\text{a}}$ is the flux with the soil surface described in Eq. (2.27), $Q_{\text{heat}-\text{a}}$ is the convective flux with the heating pipes, Q_{ven} is the heat lost by natural ventilation and the heat lost by infiltration losses, $Q_{\text{trp,cr}}$ is the latent heat effect of the crop transpiration, $Q_{\text{evp,p}}$ is the latent heat effect of evaporation in the pools (in those cases in which there are water reservoirs inside the greenhouse for the Nutrient Films Technique (NFT) irrigation system), and $c_{\text{ter}} = c_{\text{sph,a}} c_{\text{den,a}} (c_{\text{vol,g}}/c_{\text{area,ss}})$ is the product of specific heat of air, air density, and effective height of the greenhouse (greenhouse volume/soil surface area).

Heat fluxes with the heating systems. Based on the heating system facilities, the used model must be different. In the case of heating pipes, heat transfer is produced by heat convective fluxes with the pipes (see Sect. 3.1.2.2 for details). It is calculated considering that the hot water temperature is similar to the temperature of the external surface of the pipes, neglecting the effect of the convective flux between the hot water with the internal surface of the heating pipes and the conductive flux of the pipes. This term is given by the following equation:

$$Q_{\text{cnv,heat}-\text{a}} = V_{\text{cnv,heat}-\text{a}} \frac{c_{\text{area,heat}}}{c_{\text{area,ss}}} \left(U_{\text{T,heat}} - X_{\text{T,a}} \right) \tag{2.36}$$

where $c_{\text{area,heat}}$ is the heat pipe surface, $U_{\text{T,heat}}$ is the water temperature in the heating pipes and $V_{\text{cnv,heat}-\text{a}}$ is the heating convective heat transfer coefficient calculated in the same way as the rest of convective coefficients:

$$V_{\text{cnv,heat}-\text{a}} = c_{\text{cnv,heat}-\text{a1}} \left| \frac{U_{\text{T,heat}} - X_{\text{T,a}}}{c_{\text{cl,heat}}} \right|^{c_{\text{cnv,heat}-\text{a2}}} + \left[V_{\text{ws,a}} \exp \left(c_{\text{cnv,heat}-\text{a4}} D_{\text{LAI}} \right) \right]^{c_{\text{cnv,heat}-\text{a5}}} \tag{2.37}$$

$c_{\text{cnv,heat}-\text{ax}}$ being empirical parameters that have to be estimated and $c_{\text{cl,heat}}$ is the characteristic length of the heating system (in this case the diameter of the heating pipes).

On the other hand, if the energy is supplied by an air heating system supposing the heating system to be perfectly linear with respect to the control signal $U_{\text{T,heat}}$, it can be assumed that

$$Q_{\text{cnv,heat}-\text{a}} = Q_{\text{heat,en}} c_{\text{heat,ef}} U_{\text{T,heat}} \tag{2.38}$$

where $Q_{\text{heat,en}}$ is the nominal energy of the heating system, $c_{\text{heat,ef}}$ is its coefficient of efficiency, $Q_{\text{max}} = Q_{\text{heat,en}} c_{\text{heat,ef}}$ is the maximum energy that can be contributed by the system, and $U_{\text{T,heat}}$ is the heater's activation control signal (on/off).

Heat lost by natural ventilation. The heat lost by natural ventilation term is modeled according to ASAE standard EP406.3 (1998), [14]:

$$Q_{\text{ven}} = \frac{c_{\text{den,a}} c_{\text{sph,a}}}{c_{\text{area,ss}}} V_{\text{ven,flux}} \left(V_{\text{Texh,a}} - D_{\text{T,e}} \right) \tag{2.39}$$

where $V_{\text{ven,flux}}$ is the volumetric flow rate described in Eqs. (2.12), (2.11) or (2.15) and $V_{\text{Texh,a}}$ is the exhaust air temperature, calculated as a linear combination of external and internal air temperature [379]:

$$V_{\text{Texh,a}} = V_{\text{ven,reg}} X_{\text{T,a}} + \left(1 - V_{\text{ven,reg}} \right) D_{\text{T,e}} \tag{2.40}$$

where $V_{\text{ven,reg}}$ is the ventilation regime coefficient. $V_{\text{vent,reg}} = 1$ is a good approach for natural ventilation through windows (as the type of greenhouse modeled in this work), so Eq. (2.39) now becomes:

$$Q_{\text{ven}} = \frac{c_{\text{den,a}} c_{\text{sph,a}}}{c_{\text{area,ss}}} V_{\text{ven,flux}} \left(X_{\text{T,a}} - D_{\text{T,e}} \right) \tag{2.41}$$

This term includes the heat lost by infiltration losses, as shown in the equation of the volumetric flow rate (2.12), (2.11) or (2.15).

Latent heat effect of crop transpiration. The crop affects the greenhouse air temperature. As no measurements of the leaf area are usually available online, it is not possible to use a convective factor in the heat balance equation using it as a boundary variable. One way to model the effect of the crop on the air temperature is based on the latent heat due to transpiration of the plants described by Eq. (2.42),

$$Q_{\text{trp,cr}} = V_{\text{lt,vap}} M_{\text{trp,cr}} \tag{2.42}$$

where $M_{\text{trp,cr}}$ is the transpiration of the crop. Most transpiration estimators are based on the Penman–Monteith equation. In 1948, Penman derived an equation that combined the energy balance and the convective transport of vapor. Later, this model was adapted by Monteith to estimate actual evapotranspiration from plants [277]. This equation essentially combines the equation for heat transfer between the crop and the mass of the surrounding air. A simplified pseudo-physical transpiration model can be used based on two main variables: solar radiation (V_{rs}) arriving at a particular depth in the canopy plant, and vapor pressure deficit (VPD, V_{vpd}), [374]:

$$V_{\text{lt,vap}} M_{\text{trp,cr}} = \exp\left(-c_{\text{extsw,cr}} D_{\text{LAI}} \right) V_{\text{rs}} c_{\text{rs}} + V_{\text{vpd}} D_{\text{LAI}} c_{\text{vpd}} \tag{2.43}$$

where $c_{\text{extsw,cr}}$ is the light extinction coefficient for crops (it is related to the leaf inclination angle and the leaf arrangement with regard to the LAI, and provides an indication of the plant's efficiency on intercepting solar radiation). The coefficient c_{rs} is constant with appropriate dimension dependent on the crop. To obtain more reliable results, the parameter c_{vpd} is obtained for diurnal (c_{vpd_d}) and nocturnal (c_{vpd_n}) periods through calibration.

On the other hand, various authors have obtained new formulations without satisfactory results for various crops. In the case treated in this book, the crop is tomato,

so a specific transpiration model for this crop can be used, like the proposal of Stanghellini [419] also based on the Penman–Monteith equation:

$$M_{\mathrm{trp,cr}} = \frac{1}{V_{\mathrm{r,trp}}}\left(V_{\mathrm{hsat,a}} + \frac{1}{c_{\mathrm{den,a}}}\frac{V_{\mathrm{ssvp}}}{c_{\mathrm{psyco}}}\frac{V_{\mathrm{r,bl}}}{2D_{\mathrm{LAI}}}\frac{V_{\mathrm{rn,cr}}}{V_{\mathrm{lt,vap}}} - X_{\mathrm{H_{a,a}}}\right) \tag{2.44}$$

where $V_{\mathrm{hsat,a}}$ is the water concentration of the air at saturation (calculated at air temperature), c_{psyco} is the thermodynamic psychometric constant, V_{ssvp} is the slope of the saturated vapor pressure curve (calculated using the air temperature), $V_{\mathrm{rn,cr}}$ is the net radiation available to the canopy (calculated on the basis of solar radiation), and $V_{\mathrm{r,trp}}$ is a transpiration resistance described by Eq. (2.45),

$$V_{\mathrm{r,trp}} = \frac{1}{2D_{\mathrm{LAI}}}\left[\left(1 + \frac{V_{\mathrm{ssvp}}}{c_{\mathrm{psyco}}}\right)V_{\mathrm{r,bl}} + V_{\mathrm{r,s}}\right] \tag{2.45}$$

where $V_{\mathrm{r,bl}}$ is the boundary layer resistance and $V_{\mathrm{r,s}}$ is the stomatal resistance. $V_{\mathrm{r,bl}}$ depends on the aerodynamic regime that prevails in the greenhouse. In [63], the buoyancy effect is neglected when compared with the wind effect, so this resistance can be expressed with respect to the average inside air speed using Eq. (2.46),

$$V_{\mathrm{r,bl}} = 220\frac{c_{\mathrm{cl,cr}}^{0.2}}{V_{\mathrm{ws,a}}^{0.8}} \tag{2.46}$$

where $c_{\mathrm{cl,cr}}$ is the characteristic length of the crop leaf. $V_{\mathrm{r,s}}$ depends on the global radiation on the crop, the greenhouse humidity, and the crop temperature [422]. For greenhouse tomato crops, the effect of the global radiation is the most important, so it can be calculated using the approach in [63]:

$$V_{\mathrm{r,s}} = 200\left(1 + \frac{1}{\exp(0.05V_{\mathrm{rs,cr}} - 50)}\right) \tag{2.47}$$

Latent heat effect of the evaporation in the pools. This is not a typical process, but may appear if the cultivation method is NFT [151]. The greenhouse contains nonisolated pools to recycle the fertilized water to maintain the continuous water flow. The evaporation of the water of the pools affects the greenhouse climate. In the same way, the transpiration of the crop is included in the balances, a factor has been added to the latent heat term:

$$Q_{\mathrm{evp,p}} = V_{\mathrm{lt,vap}}M_{\mathrm{evp,p}} \tag{2.48}$$

where $M_{\mathrm{evp,p}}$ is the evaporation flux from the pools. The evaporation from an open water surface is produced by two main factors: The energy to provide the vaporization latent heat (solar radiation) and the capacity to move the water vapor out of the evaporation surface due to wind speed and the air humidity on the surface. Evaporation

can be calculated by mixing the aerodynamic method based on the vapor pressure deficit and the energy method based on the energy balance [90]. This mixed method is adequate for small surfaces with known climate conditions and so, the following equation is used:

$$M_{\mathrm{evp,p}} = \frac{V_{\mathrm{ssvp}}}{V_{\mathrm{ssvp}} + c_{\mathrm{psyco}}} c_{\mathrm{evp,1}} V_{\mathrm{rn,ss}} + \frac{c_{\mathrm{psyco}}}{V_{\mathrm{ssvp}} + c_{\mathrm{psyco}}} c_{\mathrm{evp,2}} V_{\mathrm{vpd,a}} \tag{2.49}$$

where $c_{\mathrm{evp,1}}$ is a factor to calibrate the effect of the net radiation on the soil surface and $c_{\mathrm{evp,2}}$ is a factor to calibrate the effect of the air vapor pressure deficit, $V_{\mathrm{vpd,a}}$ calculated as

$$V_{\mathrm{vpd,a}} = V_{\mathrm{vpsat,a}} \left(1 - \frac{X_{\mathrm{H_r,a}}}{100}\right) \tag{2.50}$$

where $V_{\mathrm{vpsat,a}}$ is the saturation vapor pressure calculated as an exponential function of the internal air temperature and $X_{\mathrm{H_r,a}}$ is the relative humidity calculated in the basis of the absolute humidity, $X_{\mathrm{H_{a,a}}}$ (see Eq. (2.52)), using the following expression:

$$X_{\mathrm{H_r,a}} = \frac{c_{\mathrm{den,a}}}{0.00217} \left(\frac{X_{\mathrm{H_{a,a}}} X_{\mathrm{T,a}}}{V_{\mathrm{vpsat,a}}}\right) \tag{2.51}$$

2.1.1.7 Water Mass Transfer Fluxes with the Internal Air

A model of absolute humidity (water vapor content of the greenhouse air) is based on a water vapor mass balance equation. As Fig. 2.2d shows the main sources of vapor in a greenhouse are crop transpiration, evaporation of the soil surface and pools, and water influx by fogging or cooling. The vapor outflow takes place through condensation on the internal side of the cover, ventilation, and vapor lost by infiltration losses. As artificial water influxes (cooling, fogging, etc.) are not installed in greenhouses in which the experiments were carried out, the mean water vapor content of the greenhouse air, $X_{\mathrm{H_{a,a}}}$, (absolute humidity) is modeled using the water mass balance equation given by Eq. (2.52),

$$\frac{c_{\mathrm{vol,g}}}{c_{\mathrm{area,s}}} c_{\mathrm{den,a}} \frac{\mathrm{d}X_{\mathrm{H_{a,a}}}}{\mathrm{d}t} = M_{\mathrm{trp,cr}} + M_{\mathrm{evp,p}} + M_{\mathrm{evp,ss}} - M_{\mathrm{cd,cv}} - M_{\mathrm{ven,a-e}} \tag{2.52}$$

where $M_{\mathrm{trp,cr}}$ is the crop transpiration flux described in Eq. (2.44), $M_{\mathrm{evp,p}}$ is the evaporation flux from the reservoirs described in Eq. (2.49), $M_{\mathrm{evp,ss}}$ is the mass evaporation flux from the soil surface described in Eq. (2.31), $M_{\mathrm{cd,cv}}$ is the condensation flux from the cover described in Eq. (2.22) and $M_{\mathrm{ven,a-e}}$ is the outflow by natural ventilation described by the following equation, where the volumetric flow rate, $V_{\mathrm{ven,flux}}$, is described in Eqs. (2.12), (2.11) or (2.15):

$$M_{\text{ven,a-e}} = \frac{c_{\text{den,a}}}{c_{\text{area,ss}}} V_{\text{ven,flux}} \left(X_{\text{H}_{\text{a,a}}} - D_{\text{H}_{\text{a}},\text{e}} \right) + M_{\text{loss,a-e}} \tag{2.53}$$

where $M_{\text{loss,a-e}}$ are infiltration losses.

2.1.1.8 Model Implementation

The designed greenhouse climate model is composed of five ODEs related to the main greenhouse climate variables (temperature and humidity of internal air, cover temperature, soil surface temperature, and first soil layer temperature) and 49 algebraic equations including the PAR radiation onto the canopy. This model is divided hierarchically using a top-down approach from a high level that includes all the submodels to the lower level where each physical process is modeled [355]. The advantages of using this hierarchical division are:

- Each submodel can be studied independently, simplifying the problem of parameter calibration.
- A new state variable submodel can be easily added, such as the crop temperature or CO_2 concentration, programming the new balance equations and adding or eliminating physical effects in the determined submodels.
- A submodel can be added or eliminated depending on the installed actuators. If a new actuator is installed (e.g., cooling), it can be modeled and added to humidity and temperature submodels easily.
- A model of a physical process can be substituted when a better model is available by changing the corresponding submodel.
- A submodel can be substituted by its real measurements when these are available, thus reducing the uncertainties because the number of variables to estimate is smaller.
- Each submodel can act separately as a "soft sensor," providing an estimate of unmeasured state variables (e.g., cover temperature) based on other measured variables.

The input/output scheme of the model is shown in Fig. 2.5 and is divided into the following submodels:

A. Temperature submodel

A.1. Cover temperature submodel
A.1.1. Cover solar radiation absorption submodel
A.1.2. Cover internal convective flux submodel
A.1.3. Cover external convective flux submodel
A.1.4. Cover condensation flux submodel
A.1.5. Cover thermal radiation absorption submodel
A.2. Soil Surface temperature submodel
A.2.1. Soil surface solar radiation absorption submodel
A.2.2. Soil surface convective flux submodel

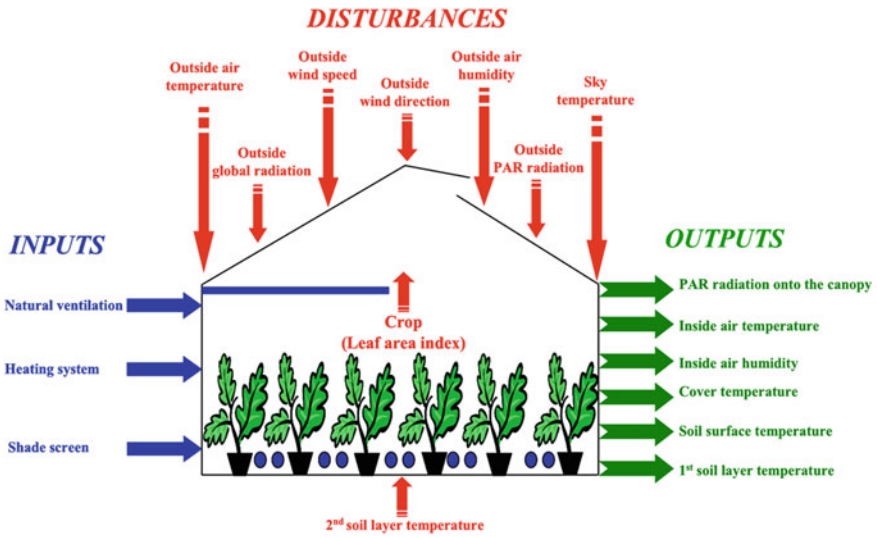

Fig. 2.5 Input–output scheme

A.2.3. Soil surface conduction to first layer submodel
A.2.4. Soil surface evaporation flux submodel
A.2.5. Soil surface thermal radiation absorption submodel
A.3. First soil layer temperature submodel
A.3.1. First soil layer conduction to soil surface submodel
A.3.2. First soil layer conduction to second layer submodel
A.4. Internal air temperature submodel
A.4.1. Air convective flux with cover submodel
A.4.2. Air convective flux with soil surface submodel
A.4.3. Air convective flux with heating submodel
A.4.4. Heat loss by ventilation submodel
A.4.5. Crop transpiration flux submodel
A.4.6. Pool evaporation flux submodel

B. Internal air humidity submodel

B.1. Cover condensation flux submodel
B.2. Soil surface evaporation flux submodel
B.3. Crop transpiration flux submodel
B.4. Pool evaporation flux submodel
B.5. Water vapor lost by ventilation submodel

C. Greenhouse PAR radiation submodel

In order to implement the model, two paradigms can be used:

- A block-based modeling and simulation approach using Simulink [268] running on Matlab. Matlab is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment [267]. Simulink is an interactive system for modeling, simulating, and analyzing linear and nonlinear dynamical models (continuous, sampled, or hybrid systems). It is a graphical mouse-driven program that allows the user to model a system by drawing a block diagram on the screen and manipulating it dynamically. Simulink includes a comprehensive block library of sinks, sources, linear and nonlinear components, and connectors, so that the user can build the model using these blocks and connecting them adequately. It is possible to add new customized blocks. Each Simulink block is composed of an input vector and output vector, and a state vector relating inputs to outputs. The main advantage of this tool is that it is not necessary to write a program as happens with other simulations tools. The initialization of the model is performed by a designed Matlab program that loads in the workspace of Matlab the greenhouse structure data (surface, volume, etc.), the characteristic of the materials used in the greenhouse (cover, soil, etc.), the features of the actuator systems (length and width of the vents, diameter of the heating tubes, etc.), universal physical constants (psicometric constant, etc.), values of the coefficients involved in the physical processes (convective and conduction coefficients, etc.), crop data (density of plants, extinction coefficient, etc.), and the initial values of state, output, characteristic, and disturbance variables. Furthermore, it reads the values of the available external variables contained in data files (note that the model could also be used for online estimation of state variables as typical sampling time is enough for their calculation, which could be included, for instance, in predictive control schemes or production optimization arquitectures). The way in which the Matlab program has been developed simplifies the use of the developed model for new greenhouse structures or new external data inputs. The greenhouse climate model has been divided into several submodels hierarchically organized in five levels:

1. System level. It consists of two blocks (climate model and crop model). The inputs (control and disturbances) and the outputs are included, as well as the relations between the systems that constitute the compound model.
2. Variable type level. It corresponds to climate variables, consisting of three models: PAR radiation, temperature, and humidity.
3. Variable level. Some climate type variables can be defined by some variables. The temperature level is divided into four submodels: Cover, soil surface, first soil layer, and greenhouse air temperatures.
4. Process level. It is formed by the submodels of physical processes involved in the models of the variables.
5. Implementation level. It corresponds with Simulink code to implement the process models of the upper level.

On the other hand, the simulation of this model involves the numerical integration of five ordinary differential equations. Simulink provides a number of solvers for the integration of such equations. Due to the diversity of dynamic system behaviors, some solvers may be more efficient than others when solving a particular problem. In the case treated in this paper, the Gears methods are used, as the greenhouse climate is a stiff problem (the system has slow and fast dynamics and these at last reach a steady state).

- An object-oriented modeling proposal using Modelica as a declarative and equation-based language for modeling multidomain physical systems [116, 117]. One natural method for physical systems modeling is to decompose the whole system in subsystems interconnected by means of their interfaces. These subsystems could decompose themselves in other interconnected subsubsystems and so on. Each subsystem is modeled using conservation laws (energy, mass, momentum, etc.) and constitutive equations in terms of differential and algebraic equations (DAE). This methodology promotes greatly building reusable models. This paradigm is different from the block-oriented modeling, presenting some advantages as the causality management. To develop the model of the compound greenhouse climate model using Modelica, the OMT (Object Modeling Technique) methodology, proposed by Rumbaugh et al. [367], is used. This tech nique proposes a formal graph showing the relations (association, aggregation, and generalization) between the different objects that constitute the systems and their properties and attributes. Three general classes are defined [355, 363]:

  `Crop_model` class. It represents the LAI (modeled or measured) of a tomato crop.

  `Greenhouse_class`. This class describes the greenhouse where the simulation test is designed. Its attributes are the parameters of the different elements constituting the greenhouse. The main advantage of this design is the possibility of changing or adding a physical element (i.e., actuators) easily. These classes are described by their own name:

  `Structure`. Type and dimensions of the greenhouse structure.

  `Ventilation`. Type and dimensions of the ventilation.

  `Heating`. Type and parameters of the heating system.

  `Soil_surface`. Type of material of the soil surface.

  `First_layer_soil`. Type of material of the first layer soil.

  `Second_layer_soil`. Type of material of the second layer soil.

  `Cover`. Type of cover material.

  `Greenhouse_model` class. It represents the different models that describe the greenhouse climate variables. It is related to the `Greenhouse_class` to obtain the parameters of the greenhouse where the simulation experiences are performed. Furthermore, this class is related with the `Crop_model` class to model the effects of the plants on the climate. It is constituted by an aggregation relation of the following classes:

  `Temperature_model`. Class of the different models of internal air temperature.

`Humidity_model`. Class of the different models of internal air humidity.
`Cover_model`. Class of the different models of cover temperature.
`Soil_model`. It is formed by two subclasses:
`Soil_surface_model`. Class of the different models of soil surface temperature.
`Soil_layer_model`. Class of the different models of the soil layer temperature.

The compound model is defined by five ordinary differential equations and 59 algebraic equations. This equation system is solved using the DASSL algorithm [65] because the simulation computational time was the smallest and it is very efficient to solve stiff systems.

The use of modeling environments as Simulink or Dymola/Modelica and systematic procedures for decomposing the complete model in submodels, which can be independently validated, has shown to facilitate the implementation of the compound model (as an integration of the single submodels) and its extension to other types of greenhouses. The choice of a simulation paradigm and implementation tool depends on the skill and ability of the user to implement the models and especially their preferences on the working methodology of each.

2.1.1.9 Model Calibration

Due to the large set of unknown parameters (more than 30), it is difficult to obtain their values using a unique search technique with the compound model. The solution consists in performing single experimental tests for each one of the involved processes to estimate their parameters in a similar way as the experiences carried out by Bot [57]. These experiments are not easy to perform, and some of them are expensive and present a long duration. On the other hand, the input/output meteorological and actuator status data are often at hand in a typical greenhouse installation, so it would be desirable to use only these data to calibrate the greenhouse climate model, without losing the physical meaning of the processes involved in the balance equations. This problem can be simplified considering the following facts:

- Data of the different climate variables to model, the disturbances and the actuators status are measured, so the problem has been divided into some submodels calibration processes (air humidity and cover, air, soil surface, and first soil layer temperature).
- Some of the involved physical processes in the balance equations are not coupled or they have no influence in determined time lapses of a day (e.g., the solar absorbance during the night or the crop presence), so all the parameters of a single submodel do not have to be estimated simultaneously.
- Some of the involved physical processes are modeled in different forms based on determined situations (as the convection process between the internal side of the cover and the greenhouse air in which the parameters of the convection

coefficient are different depending on laminar or turbulent regimes). So, the calibration process can be divided for each of these situations.
- In order to estimate the parameters related to the actuation systems, some guided test (mainly step response and impulse response ones) can be performed at the real greenhouse.

Based on these considerations, a methodology to calibrate the compound model was proposed by Rodríguez [355]. In what follows, the step sequence that has to be carried out to calibrate the implemented model for any greenhouse is briefly explained based on the typical measured data in a greenhouse. In each step, the number of the estimated parameters is indicated:

1. Calibration of the climate variables with an empty greenhouse (without crop)
    a. Climate variables calibration without the effects of the actuation systems (no heating, no ventilation)
        i. Calibration of the first soil layer temperature submodel [4 parameters]
        ii. Calibration during nocturnal time intervals (without solar radiation)
        iii. Calibration of cover temperature submodel
        iv. High wind speed [1 parameter]
        v. Low wind speed [3 parameters]
        vi. Calibration of soil surface submodel [6 parameters]
        vii. Calibration during diurnal time intervals (with solar radiation)
        viii. Calibration of cover temperature submodel [3 parameters]
        ix. Calibration of soil surface temperature submodel [3 parameters]
        x. Calibration of internal air humidity submodel [2 parameters]
    a. Calibration of the parameters related to natural ventilation (without heating) [2 parameters]
    b. Calibration of the parameters related to heating system (without vents) [2 parameters for pipe heating systems or 1 parameter for air heaters]
2. Calibration of the climate variables with crop
    a. Calibration of the long wave parameters in the cover temperature submodel [1 parameter]
    b. Calibration of the long wave parameters in the soil surface temperature submodel [2 parameters]
    c. Calibration of the parameters related with the crop transpiraton process [4 parameters]
3. Calibration of the PAR radiation model [1 parameter]

The largest number of parameters to estimate simultaneously is six in the processes of soil surface calibration in nocturnal time intervals with an empty greenhouse. The use of an adequate parameter search technique can help to solve this problem. In order to obtain the unknown parameters in the equations described in Sect. 2.1.1.2, a large set of input/output data obtained at the real greenhouses are used in such a

way that the values of the parameters are obtained by minimizing a least squares criterion:

$$J = \|\mathbf{X}_{\text{real}} - \mathbf{X}_{\text{sim}}\|^2 = \sum_{i=1}^{N} (X_{\text{real}}(i) - X_{\text{sim}}(i))^2 \tag{2.54}$$

where $\mathbf{X}_{\text{real}} = (X_{\text{real}}(1), \ldots, X_{\text{real}}(N))$ is a set of N real measurements of the variables to estimate and $\mathbf{X}_{\text{sim}} = (X_{\text{sim}}(1), \ldots, X_{\text{sim}}(N))$ are the values of the variables calculated by the implemented model. The used parameter search technique is divided into two phases.

In a first phase, the submodels were calibrated independently using a direct sequential search [330], consisting in an iterative method incrementing the values of the parameters between upper and lower limits (wide margins) with a determined step until a n-tuple of parameters that minimizes the least square criterion is found. The initial upper and lower bounds were obtained from physical properties and from values found in the literature. The search can be improved by decreasing the limits and the sequential increment step. The main disadvantage of this type of techniques is the high computational cost because it must evaluate all the values of the search space. So, it is used to obtain only approximated values of the model parameters reducing the search space.

```
Begin;
time=0;
generate initial population, P0;
evaluate P0;
while not finish-condition do
    begin;
    time=time+1;
    select potential solutions Mtime from Ptime-1;
    alter Mtime using genetic operators;
    create new population Ptime from Mtime;
    evaluate Ptime;
end
End.
```

Algorithm 1: Parameter search technique.

In a second phase, genetic algorithms (GAs) were used as heuristic search technique to refine the obtained values of the model parameters in the first phase. GAs are globally oriented in searching and thus potentially useful in solving optimization problems in which the objective functions response contain multiple optima and other irregularities [172, 329, 463, 475]. Empirical studies have demonstrated that GAs have been successfully applied to several types of problems, including function optimization or model fitting [369]. GAs differ from the iterative search in that they search among a population of points and use probabilistic rather than deterministic transaction rules. These algorithms are formulated using a direct anal-

ogy with evolution processes observed in nature. GAs work simultaneously with a population of individuals (n-tuples of parameters) exploring a number of new areas in the search space in parallel, thus reducing the probability of being trapped in a local minimum [269]. As in nature, individuals in a population compete with each other for surviving, so that fitter individuals tend to progress into new generations, while the poor ones usually die out. This process is described in Algorithm 1.

The initial population is randomly generated within certain boundaries. The determination of these boundaries is a difficult problem. In the case treated in this section, these limits were determined by the study performed with the sequential search phase. In order to evaluate the population, the simulation is run for each individual (set of all model parameters to estimate), and a numerical value is assigned to each member of the population (possible set of model parameters) using the least squares criterion given in Eq. (2.54). All the individuals in the population are evaluated and their fitness are used as the basis of the selection. A common selection approach assigns a probability of selection, $P(j)$, to each individual j based on its fitness value. A series of N random numbers is generated and compared against the cumulative probability of the population:

$$C(i) = \sum_{j=1}^{i} P(j) \tag{2.55}$$

The appropriate individual, i, is selected to belong to the new population if $C(i-1) < U(0, 1) < C(i)$ where $U(0, 1)$ is a uniform distribution. In [190] different methods to assign probabilities to individuals are proposed, such as the roulette wheel and ranking methods. A normalized geometric ranking method has been used in this application. It assigns a probability $P(i)$ based on the rank of solution i when all solutions are sorted. The method defines $P(i)$ for each individual using the following equation:

$$P(i) = \frac{P_{\text{best}}}{1 - (1 - P_{\text{best}})^{P_s}} (1 - P_{\text{best}})^{\text{rank(i)}-1} \tag{2.56}$$

where P_{best} is the probability of selecting the best individual, P_s is the population size and rank(i) is the rank of the individual where 1 is the best. In order to alter the selected individuals to generate the new population, GAs used two basic types of operators:

- Crossover. This operator takes two individuals and produces two new individuals exchanging genetic information in pairs or larger group between the parents. There are some crossover operators and in this application, the real value simple crossover has been used. Let $\mathbf{X}$ and $\mathbf{Y}$ two m-dimensional row vectors of floats denoting individuals from the population. This operator generates a random number n from a uniform distribution $U(0, 1)$ and creates two new individuals $\acute{\mathbf{X}}$ and $\acute{\mathbf{Y}}$ based on the following equations:

$$\begin{aligned} \acute{\mathbf{X}} &= n\mathbf{X} + (1-n)\mathbf{Y} \\ \acute{\mathbf{Y}} &= (1-n)\mathbf{X} + n\mathbf{Y} \end{aligned} \tag{2.57}$$

- Mutation. This operator alters one individual to produce a single new solution. A uniform mutation algorithm has been used that selects randomly one variable j, and sets it equal to a uniform random number $U(a_i, b_i)$ where $\underline{a}(i)$ and $\overline{a}(i)$ are the lower and upper limit of the interval of variation of the parameter at parent chromosome position, that is:

$$\acute{X}(i) = \begin{cases} U(\underline{a}(i), \overline{a}(i)) & \text{if } i = j \\ X(i) & \text{if } i \neq j \end{cases} \tag{2.58}$$

Table 2.1 and Fig. 2.6 (mean absolute error, maximum absolute error and standard deviation) show that the estimation of the coefficients using GAs methods is better than the iterative search for the complete model. Furthermore, this technique is more efficient in time. Although it is necessary to indicate that the search space of the

Table 2.1 Comparison between real temperature (°C) and relative humidity (%) and the estimation of the compound model using GAs (GA) and direct sequential search (Sec) during August 2000 in Araba greenhouse

	Air temperature		Cover temperature		Soil surface temperature		First soil layer temperature		Air relative humidity	
	Sec	GA	Sec	GA	Sec	GA	Sec	GA	Sec	GA
Mean absolute error	1.13	0.93	0.89	0.79	0.74	0.64	0.34	0.32	4.29	3.92
Maximum absolute error	5.15	4.63	5.17	4.81	4.66	4.65	1.36	1.26	29.47	24.32
Standard deviation	0.88	0.65	0.79	0.64	0.63	0.50	0.28	0.26	3.76	3.69

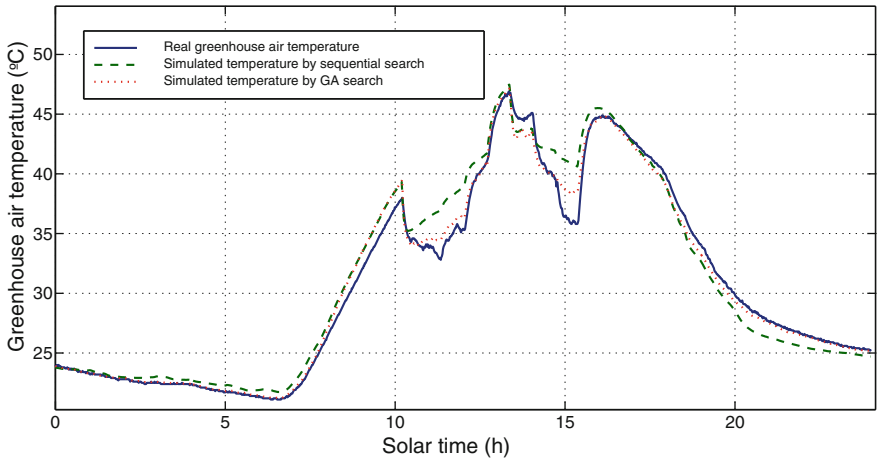

Fig. 2.6 Comparison between real temperature and complete model simulated temperatures of greenhouse air using GAs and sequential search

initial population was reduced because it was deduced of the study carried out with the sequential search of the simple submodels.

In order to calibrate the parameters of the greenhouse climate models, several tests were performed in the greenhouses: In summer season (without crop and with guided experiences using natural ventilation) from June to August, and others in winter season (with tomato crop and guided experiences using the heating system) from December to March. Data of 15 days with 1-min sample time (21,600 real measurements) were used in the calibration phase in each season. In order to calibrate the PAR radiation, modeled by an algebraic equation, data of a month without whitening were used (January) to verify the data provided by the manufacturers of the cover material and the shade screen. The values are slightly corrected because they lose the original properties along time. This variation of the parameters is not accounted for by the model because the chemical equations that describe the degradation of the physical characteristics are not known. In any case, the degradation process takes place slowly, so it is logical to suppose that these parameters are constant during a simulation experiment (during one season at the most). It is obviously necessary to calibrate these values along time. The used data calibrating the effect on the transmission coefficient of the cover when it is whitened correspond to May and June. The submodels are independently calibrated because all the needed input/output data are measured. The calibration process is similar for any greenhouse, so only the obtained results for Araba greenhouse (see Appen. A) are shown in this section for lack of space.

Some of the results of the calibration processes are shown in Fig. 2.7 and Table 2.2, where a comparison between real measurements and those obtained by separate simulations are shown using data of August for all the variables and data of January for air temperature and humidity (which are the main variables). These results are different depending on the models defined by the known state variables:

- Configuration 1. Simple submodel with full measurements of the other state and exogenous (external) variables. The model only estimates the air temperature.
- Configuration 2. Using real data of cover and soil temperatures and exogenous variables. The model estimates the air greenhouse temperature and humidity.
- Configuration 3. Using real data of soil temperatures and exogenous variables. The model has to estimate the greenhouse air temperature and humidity, and the cover temperature.
- Configuration 4. Using real data of humidity and exogenous variables. The model has to estimate the temperature of the first soil layer, the soil surface, the cover, and the greenhouse air.
- Configuration 5. Without supplying data of humidity, cover temperature, and soil temperatures (complete model). All the state variables are simulated and only external variables are supposed to be known.

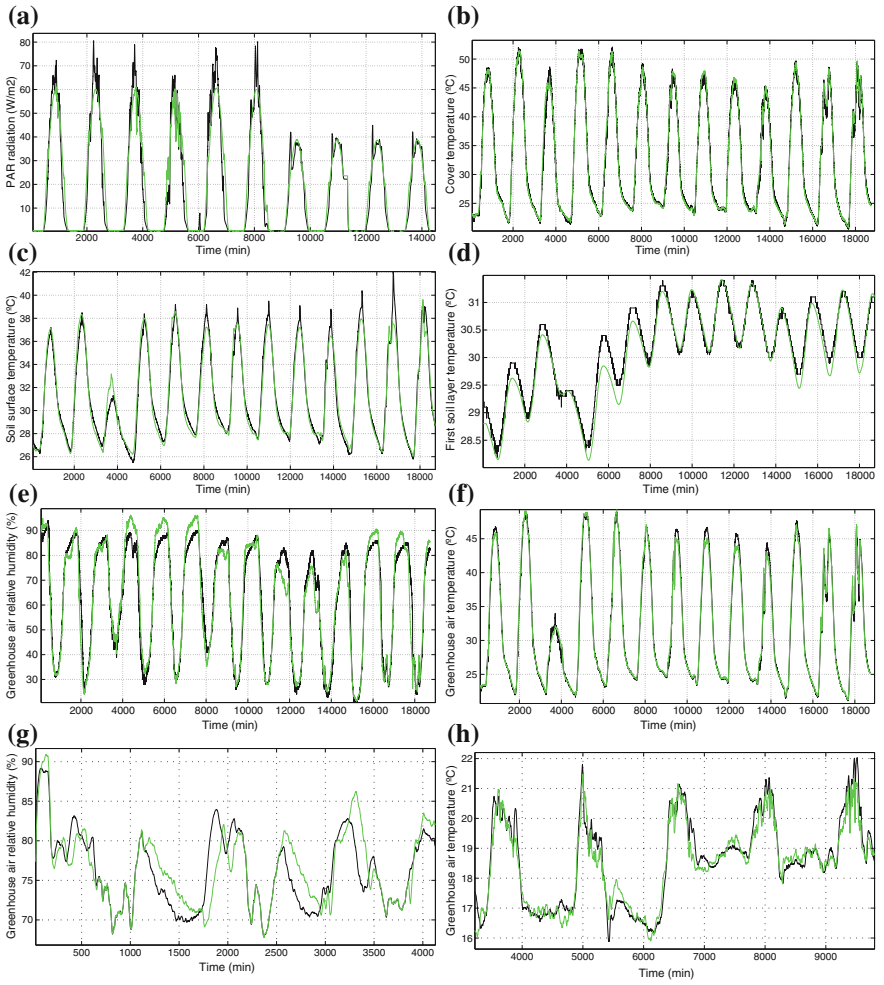

Fig. 2.7 Simulation results with the independent submodels in the calibration process. **a** PAR radiation submodel in spring. **b** Cover temperature submodel in summer. **c** Soil surface submodel in summer. **d** First soil layer submodel in summer. **e** Greenhouse humidity submodel in summer. **f** Greenhouse temperature submodel in summer. **g** Greenhouse humidity submodel in winter. **h** Greenhouse temperature submodel in winter

Table 2.3 provides the results of the temperature estimated by the five configurations, in terms of maximum, mean and standard deviation absolute errors when compared with real data. As more variables are required to be modeled, larger errors are obtained, as shown in Fig. 2.8. This is due to the fact that the uncertainties in the modeled processes increase the numerical errors, which are greater because it is necessary to solve a larger number of equations. This result was predictable, although the behavior of the model can be considered adequate in every configuration because

Table 2.2 Comparative results of the estimation of the different climate variables in the calibration process

	Summer				Winter		
	Air temp.	Air relative humidity	Cover temp.	Soil surface temp.	First soil layer temp.	Air temp.	Air relative humidity
Variation interval	21.1–49.0 (27.9 °C)	21–94 (73 %)	20.55–52.1 (31.55 °C)	25.5–42 (16.5 °C)	28.19–31.4 (5.9 °C)	11.5–25.5 (14 °C)	47.9–100 (50.3 %)
Mean	0.51	3.96	0.52	0.68	0.25	0.52	2.53
Maximum	2.81	24.32	3.38	4.12	0.79	2.06	17.19
Standard deviation	0.52	3.75	0.53	0.44	0.17	0.48	2.39

Table 2.3 Comparative results of the estimation of the greenhouse air temperature using different configurations

	Configuration 1	Configuration 2	Configuration 3	Configuration 4	Configuration 5
Mean	0.51	0.52	0.61	0.93	0.95
Maximum	2.81	2.83	3.12	4.63	4.73
Standard deviation	0.52	0.53	0.59	0.65	0.66

the mean of the absolute errors is not greater than 4 % within the variation interval of the greenhouse air temperature. As the results show, the model calibration process has been successfully performed.

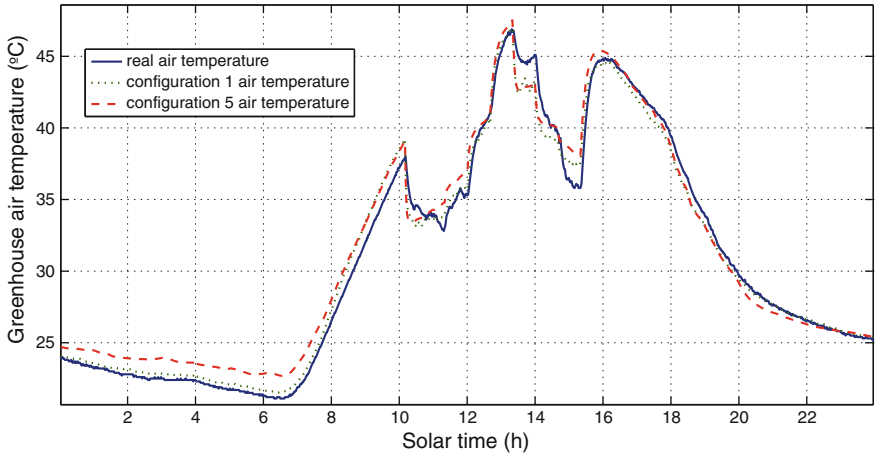

Fig. 2.8 Simulation results with the independent submodels in the calibration process

2.1.1.10 Sensitivity Analysis

A sensitivity analysis based on the variation of the optimal obtained parameters has been carried out to study the robustness of the model formulation. The numerical methods proposed by Cameron [75] have been used, consisting in the calculation of the cost function described in Eq. (2.54) for 21 values of each one of the model parameters, in a variation interval of $\pm 10\,\%$ with respect to their optimal values. As an example, Fig. 2.9 shows the greenhouse air temperature sensitivity analysis for summer time (without crop) and during winter time (with crop and heating system). A zoom of the obtained results shows that the curves corresponding to the variations of the cost function are not symmetrical with respect to the minimum. This is mainly due to the nonlinear dependence of the models with the parameters. Even so, it is observed that around $\pm 5\,\%$ of the parameters optimal value, the linearization hypothesis is valid and so the selection of a quadratic cost function can be considered correct. The conclusions of the sensitivity analysis of each submodels are the following:

- Air temperature submodel. As Fig. 2.9a shows, in the case of an empty greenhouse during summer time conditions, this model is more sensible to the parameters of the convection process between the cover and the greenhouse air. This fact can be explained because the greenhouse air temperature depends on the outside climate and the cover acts like a union among them. Therefore, a small variation in this process will cause a great difference between the real value and the model estimation. Also, there are two sets of parameters providing a same variation of the cost function (Fig. 2.9b shows with more detail these two groups). These are those related to the convection with the soil surface, because the soil acts as a climate regulator, providing energy during the nocturnal periods. Figure 2.9b also

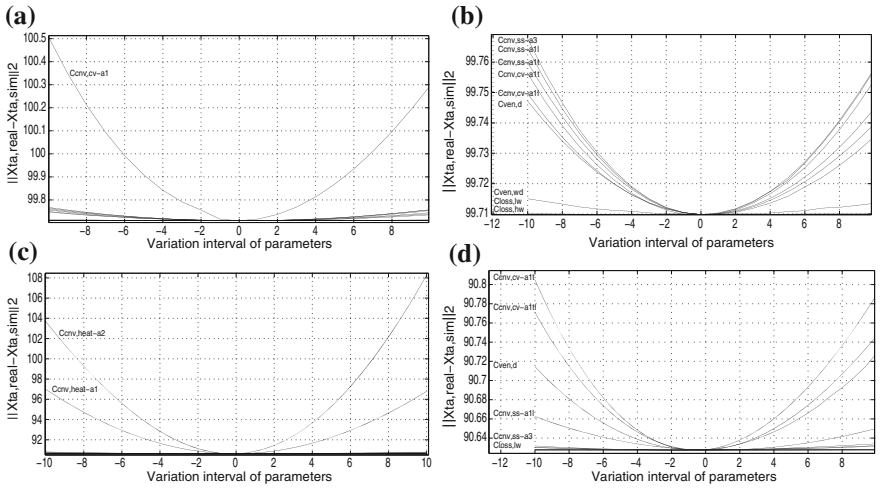

Fig. 2.9 Greenhouse air temperature sensitivity analysis. **a** Summer time complete analysis. **b** Summer time detail analysis. **c** Winter time complete analysis. **d** Winter time detail analysis

shows that the air temperature has low sensitivity to the parameters related to the ventilation process. This is not what was expected as the ventilation is the main cooling source. Some sensitivity analyses were performed for diurnal periods of 10 h where the ventilation was acting during a long time interval, observing that all the obtained results are similar. A possible cause of the obtained result is the low ventilation rate of the ventilation system installed in the analyzed greenhouse. This result should not be extrapolated to other greenhouse structures. In the case of a greenhouse with crop under winter conditions, the temperature of the air is more sensible to the convection process with the heating pipes parameters as Fig. 2.9c shows. Figure 2.9d shows a zoom of the influence of the rest of parameters. The results are similar to the previous analysis, where the model is sensible to cover convection parameters, moderately robust to changes in the parameters of soil convection and quite robust to the variation of the ventilation parameters.

- Cover temperature submodel. The cover temperature submodel is quite sensible to the parameters related to long wave radiation between the cover and the rest of the solids (soil, crop, and heating system) due to the fact that the temperature difference between the different elements is the source of the processes of heat transmission. The temperature of the heating system is very high compared with that of the other elements, reason of why its effect is larger. The heat transmission by thermal radiation depends on the temperature difference power to 4, reason of why its contribution is very important.
- First soil temperature submodel. This model is more sensible to the conduction coefficient with the soil surface as was to be expected. It is observed that the degree of sensitivity with respect to other parameters is of similar order, since the value of the cost function varies between 45 and 48, reason of why a special sensitivity to anyone of the parameters cannot be deduced.
- Soil surface temperature submodel. Some of the previous conclusions are extrapolated to the soil surface temperature submodel, which is more sensible to the variation of those parameters related to the processes of thermal radiation. In a second level, the model is more sensible to the conduction processes than to the convection processes with the inner air, due to the fact that the soil is a thermal buffer where the conduction processes are dominant.
- Humidity submodel. The sensitivity analysis of the humidity submodel has been divided into two stages. The first one corresponds to a period without crop under summer conditions, where the humidity model is more sensible to the parameters related to the evaporation process in the irrigation pools, mainly with the parameter related to the solar radiation. This is logical as under these conditions, this process is the main source of water contribution to the greenhouse air. On the other hand, it is less sensible to the parameters related to the natural ventilation as happened with the temperature submodel previously commented. In the second stage, a tomato crop with a middle-development state (LAI $= 3$) was considered, where the humidity model is more sensible to the parameters related to the processes of thermal radiation. This is due to the fact that, in this case, the main source of water contribution is the crop transpiration that directly depends on the net radiation

that reaches the canopy (related to the short wave and the thermal radiation). The sensitivity to the rest of parameters is similar to that of the first period.

2.1.1.11 Model Validation

As some state variables are not measured (cover and soil layers temperatures), only the configuration 5 of the compound climate model has been used to validate the greenhouse air temperature and humidity. Due to the fact that all the state variables are related, if two of them are validated, it can be expected that the behavior of the rest of them is adequate. In any case, the estimation of these variables provided by the model has been studied to confirm that their evolution is that expected. The experiences performed to validate the model are the following:

- Model validation with data of Araba greenhouse. After the calibration, the model for this greenhouse (described in the previous sections) with data of winter and summer seasons, the following tests were performed:
  - Evaluation of the model in the same seasons of another year: January and August.
  - Evaluation of the model in a different season of those used in the calibration process: Spring season.

  Figure 2.10 shows some results of these experiences. Analyzing these data, the validity of the developed model can be confirmed both from quantitative and qualitative viewpoints, because it follows the dynamics of the modeled variables and the errors are within acceptable intervals for this type of applications (the relative error of the absolute error average is less than 7 %). Obviously, this assertion is valid only for this greenhouse, so in order to generalize this conclusion, it was necessary to perform new validation experiences in other greenhouses with different structures, different actuators, and different control strategies as is commented in what follows.

- Model validation with data of Araba greenhouse number 3. This greenhouse is similar to Araba greenhouse except the position of the roof vents. So it was necessary to carry out the calibration of the ventilation parameters. The rest of the parameters are the same of the Araba greenhouse. In order to validate the model, some tests were carried out for three different seasons: Winter (January), spring (April), and summer (August). Table 2.4 shows some results of these experiences. The conclusions are similar to the another Araba greenhouse with a relative error of the absolute error average less than 8 %.
- Model validation with data of Inamed greenhouse. This is a hard test for the model structure as this greenhouse is different from the previous ones. After the estimation of the model parameters for Inamed structure (different to the Araba structure) using the data corresponding to the winter and summer seasons, three different experiences were performed to validate the climate model under different climate

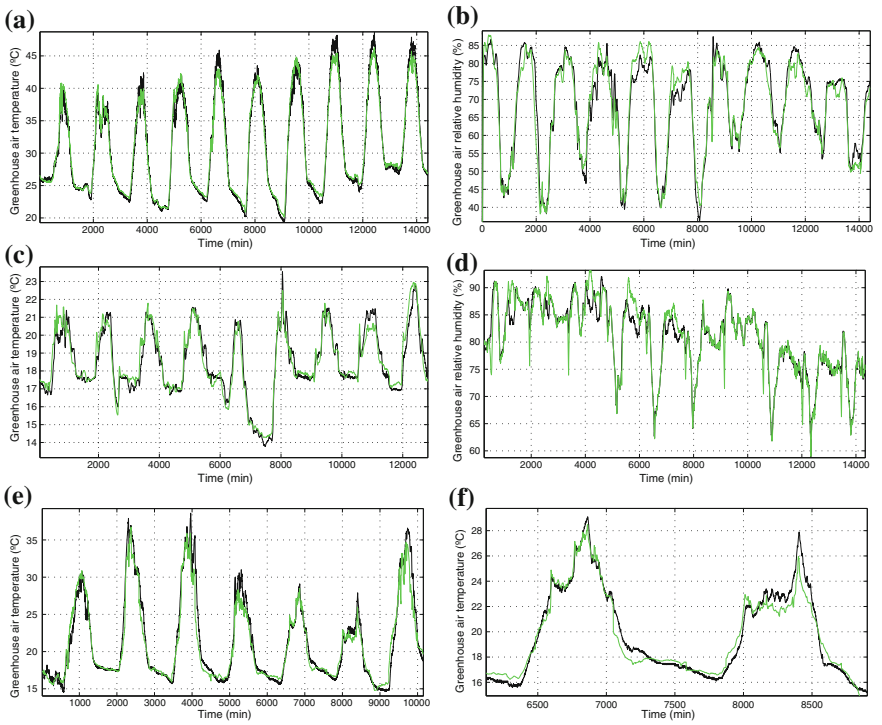

Fig. 2.10 Simulation results of Araba greenhouse in the validation process. **a** Greenhouse air temperature in summer. **b** Greenhouse air relative humidity in summer. **c** Greenhouse air temperature in winter. **d** Greenhouse air relative humidity in winter. **e** Greenhouse air temperature in spring. **f** Detail of greenhouse air temperature in spring

Table 2.4 Comparative results of the estimation of the different climate variables in Araba greenhouse number 3 in the validation process

	January		April		August	
	Temperature	Humidity	Temperature	Humidity	Temperature	Humidity
Variation interval	11.43–21.67 (10.24 °C)	45.4–99.1 (53.7 %)	11.3–27.3 (16.0 °C)	29.3–58.66 (59.36 %)	18.5–51.1 (32.6 °C)	31.42–92.21 (60.79 %)
Mean	0.56	4.11	0.58	4.54	1.12	3.62
Maximum	4.25	17.85	3.99	20.84	6.05	14.89
Standard deviation	0.52	3.99	0.58	4.09	0.94	3.43

conditions: Winter (January), spring (April), and summer (August). Figure 2.11 and Table 2.5 show some of the results, obtaining similar conclusions as those previously discussed.

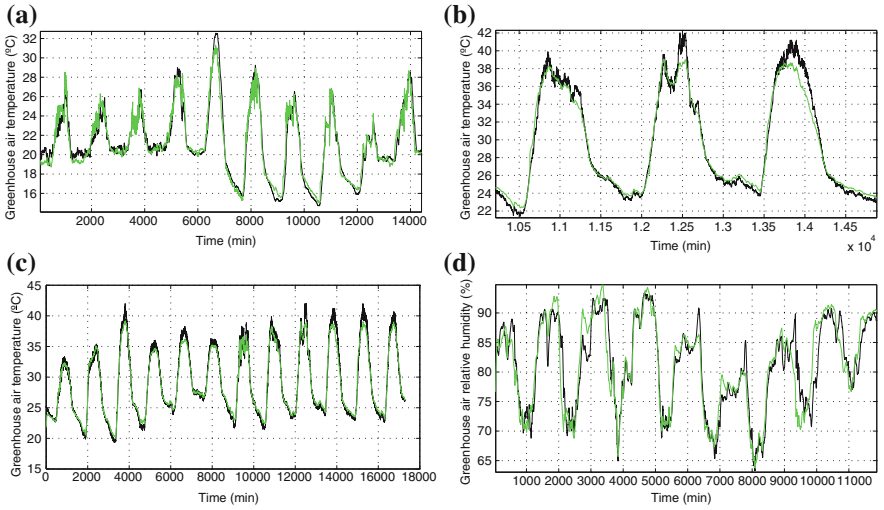

Fig. 2.11 Simulation results of Inamed greenhouse in the validation process. **a** Greenhouse air temperature in spring. **b** Detail of greenhouse air temperature in spring. **c** Greenhouse air temperature in summer. **d** Greenhouse air relative humidity in summer

Table 2.5 Comparative results of the estimation of the different climate variables in Inamed greenhouse in the validation process

	January		April		August	
	Temperature	Humidity	Temperature	Humidity	Temperature	Humidity
Variation interval	12.54–23.66 (11.12 °C)	59.4–100 (40.6 %)	14.72–32.53 (17.81 °C)	63.18–93.41 (30.23 %)	18.5–51.1 (32.6 °C)	31.42–92.21 (60.79 %)
Mean	0.48	3.26	0.63	2.11	1.12	4.01
Maximum	3.12	16.01	4.89	12.99	6.05	15.54
Standard deviation	0.43	3.17	0.55	2.19	0.94	3.97

- Model validation with data of Almería greenhouse. The same procedure was applied to this kind of greenhouse and the obtained results were similar to the Inamed case (not included for sake of space).

As it can be observed, simulations with a high degree of exigency were carried out to validate the compound model, using data of other seasons (different to those used in the calibration process) and with different greenhouse structures, obtaining adequate results that confirm the validity and performance of the model.

2.1.2 Pseudo-Physical Climate Models

2.1.2.1 General Considerations

As it has been analyzed in Sect. 2.1.1.2, the climate inside a greenhouse can be described by a system of first-order ODE which characterize the dynamics of the air, crop, soil and cover temperatures, air humidity, and CO_2 concentration. Such model is very useful for simulation and optimization purposes, but for other applications, such as climate control, simplified versions capturing the dominant dynamics of the system can be used. Several authors have proposed simple models keeping some physical sense [38, 177, 202, 384, 441, 448, 459, 460, 461]. To derive a simplified version of the model developed in Sect. 2.1.1.2, the following simplifications have to be done:

- The cover is not considered as a greenhouse element characterized by its temperature, but an interface between the inside and outside air where energy is exchanged depending on the inside-outside temperature difference. Thus, models of convection processes between the cover and the outside and inside air and the conduction process between its two surfaces are replaced.
- The crop is not considered an element and its effect on climate is modeled by transpiration and CO_2 supply or consumption due to photosynthesis and respiration. The modeling of these contributions can be done either using empirical relationships with climatic variables or detailed models developed by other authors.
- When modeling air temperature, the fundamental heat sources are the sun and the heating system, while ventilation and losses through the cover are the main heat losses. The effect on the temperature of the crop is usually taken into account, whereas the radiation through the cover is used both by the plants to perform transpiration and photosynthesis and to heat the air. The latent heat due to condensation on the cover or the evaporation in the soil surface or pool are not taken into account.
- The model of water vapor content in the air has as fundamental contributions crop transpiration and humidification systems, and ventilation as the main cause of moisture loss by exchange with the outside.
- The model of CO_2 concentration has as main inputs the artificial CO_2 supply systems and the crop respiration, and ventilation and photosynthesis as the main losses.
- Some authors include the model of the soil surface temperature in which only the energy fluxes due to convection processes with greenhouse air and conduction ones with the first soil layer (boundary condition) are taken into account. The thermal radiation processes among physical elements of the greenhouses are not considered.
- Although air is inert to radiation, most simplified models of inside air temperature include a term depending on global radiation to model air warming due to the sun. It used to be a constant factor between 0 and 1 multiplying the solar radiation transmitted through the cover.

- Heat transfer coefficients with the soil or heating pipes are considered constant and not a function of the temperature difference between the solid and the fluid or the velocity of the latter.
- The model of several physical processes such as ventilation is simplified, often using empirical relationships or considering some energy fluxes in steady state (constant), such as those from the heating system.

2.1.2.2 General Hypotheses and Simplified Model Development

The simplified pseudo-physical climate model developed in this section for control purposes is developed under the following hypotheses:

1. The state variables of the system are the inside air temperature $X_{\mathrm{T,a}}$ and humidity (absolute $X_{\mathrm{H_{a,a}}}$ and relative $X_{\mathrm{H_r,a}}$). The CO_2 concentration is not modeled because CO_2 enrichment systems are not available, but this variable is measured.
2. Three main external systems interact with the greenhouse: Outside air, soil surface, and crop.
3. The exogenous variables and disturbances acting on the system and considered as boundary conditions are the outside air temperature $D_{\mathrm{T,e}}$ and absolute humidity $D_{\mathrm{H_a,e}}$, wind speed $D_{\mathrm{ws,e}}$ and direction $D_{\mathrm{wd,e}}$, global radiation $D_{\mathrm{rs,e}}$, soil surface temperature $D_{\mathrm{T,ss}}$ and LAI as measurement of the state of the crop D_{LAI}.
4. The control inputs are the vents positions U_{ven}, the shade screen position U_{shd}, and the temperature of the water within the pipes of the heating system $U_{\mathrm{T,heat}}$ (or the heater activation control signal in the case of air heating systems).
5. A uniform homogeneous distribution of variables is considered in the air volume.
6. With respect to the processes associated with solar radiation, the following assumptions are made: Air is not inert to radiation (it absorbs and transmits radiation). Reflection effects are not considered.
7. In the heating by hot water pipes installation, water temperature is measured 1 m downstream the mixing valve, but the convection with air is done by the external surface of the tubes. The assumption is to disregard the effects of convection between hot water and inner surface of the pipes and conduction between inside and outside of the tubes, so that it is considered that the temperature of the outer surface of the pipes is equal to that of the water flowing through them.
8. The physical properties of air, such as density or specific heat, are considered constant with respect to temperature and time.

2.1.2.3 Internal Air Temperature Model

The greenhouse air temperature can be modeled using the following balance:

$$c_{\mathrm{sph,a}} c_{\mathrm{den,a}} \frac{c_{\mathrm{vol,g}}}{c_{\mathrm{area,ss}}} \frac{\mathrm{d}X_{\mathrm{T,a}}}{\mathrm{d}t} = Q_{\mathrm{sol,a}} + Q_{\mathrm{cnv,ss-a}} + Q_{\mathrm{heat-a}} + Q_{\mathrm{cnv-cnd,a-e}} - Q_{\mathrm{ven,a-e}} - Q_{\mathrm{loss,a-e}} - Q_{\mathrm{trp,cr}} \tag{2.59}$$

where $Q_{\text{sol,a}}$ is the solar radiation absorbed by the air, $Q_{\text{cnv,ss-a}}$ is the convective flux with the soil surface, $Q_{\text{heat-a}}$ is the flux with the heating pipes, $Q_{\text{cnv-cnd,a-e}}$ is the convective flux with the cover, Q_{ven} is the heat lost by natural ventilation, $Q_{\text{loss,a-e}}$ is the heat lost by infiltration losses, $Q_{\text{trp,cr}}$ is the latent heat effect of the crop transpiration, and $c_{\text{ter}} = c_{\text{sph,a}} c_{\text{den,a}} (c_{\text{vol,g}}/c_{\text{area,ss}})$ is the product of specific heat of air, air density and effective height of the greenhouse (greenhouse volume/soil surface area). These fluxes can be modeled in different ways. In the case treated in this book, the following paragraphs contain the terms used.

Solar radiation absorbed by the air. The solar radiation transmitted through the cover and reaching the crop $V_{\text{rs,cr}}$ is determined by:

$$V_{\text{rs,cr}} = V_{\text{tsw,g}} D_{\text{rs,e}} \tag{2.60}$$

where $V_{\text{tsw,g}}$ is the short wave heat transmission coefficient, which depends on the heat transmission coefficient of the cover, the whitening state, and the shading screen, as indicated by Eq. (2.5). The solar radiation absorbed by the air $Q_{\text{sol,a}}$ is given by:

$$Q_{\text{sol,a}} = c_{\text{asw,a}} V_{\text{rs,cr}} \tag{2.61}$$

where $c_{\text{asw,a}}$ is the short-wave absorption coefficient of the greenhouse air, although as the air is inert to solar radiation, it is mostly a parameter of thermal efficiency of solar energy. This coefficient must be estimated in the model calibration process.

Convective heat transfer between the soil surface and the inside air. The heat transfer between the soil surface and the inside air $Q_{\text{cnv,ss-a}}$ is a function of the temperature difference between soil surface temperature $X_{\text{T,ss}}$ and inside air temperature $X_{\text{T,a}}$, Eq. (2.62),

$$Q_{\text{cnv,ss-a}} = c_{\text{cnv,ss-a}} (X_{\text{T,ss}} - X_{\text{T,a}}) \tag{2.62}$$

where $c_{\text{cnv,ss-a}}$ is a convection coefficient considered constant and that has to be estimated.

Heat transfer by convection and conduction in the cover between the outside and the inside air. This process is considered proportional to the temperature difference between outside air temperature, $D_{\text{T,e}}$ and inside air temperature, $X_{\text{T,a}}$:

$$Q_{\text{cnd-cnv,a-e}} = c_{\text{cnd-cnv,a-e}} (X_{\text{T,a}} - D_{\text{T,e}}) \tag{2.63}$$

where $c_{\text{cnd-cnv,a-e}}$ is a thermal loss coefficient (considering convection and conduction processes) which is considered constant and is estimated empirically.

Heat transfer by the heating system. The same models described in previous sections can be used (Eqs. (2.36) and (2.38)).

Heat transfer to the outside air due to ventilation and infiltration. As mentioned above, both fluxes are modeled simultaneously as the infiltration losses process is included as a constant effect in the ventilation flux $V_{\text{ven,flux}}$, as evidenced by Eqs. (2.12), (2.11) and (2.15). Therefore, the following model is used to describe these processes:

$$Q_{\text{ven,a-e}} + Q_{\text{loss,a-e}} = \frac{c_{\text{den,a}} c_{\text{sph,a}}}{c_{\text{area,ss}}} V_{\text{ven,flux}} (X_{\text{T,a}} - D_{\text{T,e}}) \tag{2.64}$$

The ventilation flux is described by Eqs. (2.12), (2.12) and (2.15). Another option is to consider a simplified volumetric flow rate using an exponential expression of the aperture control signal (this is usual in greenhouses of Almería type [325]):

$$V_{\text{ven,flux}} = c_{\text{ven,n}} c_{\text{ven,l}} c_{\text{ven,w}} D_{\text{ws,e}} (\alpha_v U_{\text{ven}}^{\beta_v}) + V_{\text{loss}} \tag{2.65}$$

where U_{ven} in this case is the percentage or normalized aperture of the vents, $c_{\text{ven,n}}$ is the number of vents, $c_{\text{ven,l}}$ is the length of the vents, $c_{\text{ven,w}}$ is the width of the vents, and α_v and β_v are tuning parameters which, according to actual measurements, show subtle variations between leeward and windward ventilation, and V_{loss} is the leakage when the vent is closed. This is a very simplified expression as the effective opening surface should have to be used through variable $V_{\text{ven,hef}}$, as in Eqs. (2.12) and (2.11) (Eq. (2.15) if both roof and sidewall ventilation openings are considered—in that case Eq. (2.65) should include two terms accounting for both control signals), but it has demonstrated to be valid for the kind of greenhouses considered in this book [326].

Latent heat transfer by crop transpiration. The effect of crop transpiration on the inside air temperature can be modeled using Eq. (2.42), considering that the net radiation absorbed by the crop is equal to the solar radiation neglecting the effect of thermal or long-wave radiation and that the boundary layer resistance, $V_{\text{r,bl}}$, can be considered constant and equal to $200\,\text{s}\,\text{m}^{-1}$ in the range of wind speeds inside the greenhouse [420].

2.1.2.4 Internal Air Humidity Model

The greenhouse air humidity can be modeled using the following equation:

$$c_{\text{den,a}} (c_{\text{vol,g}} / c_{\text{area,ss}}) \frac{\text{d}X_{\text{H,a}}}{\text{d}t} = M_{\text{trp,cr}} - M_{\text{ven,a-e}} \tag{2.66}$$

where the main source of water vapor is crop transpiration $M_{\text{trp,cr}}$, described in Eq. (2.44), while the primary source of loss of water vapor is produced by the

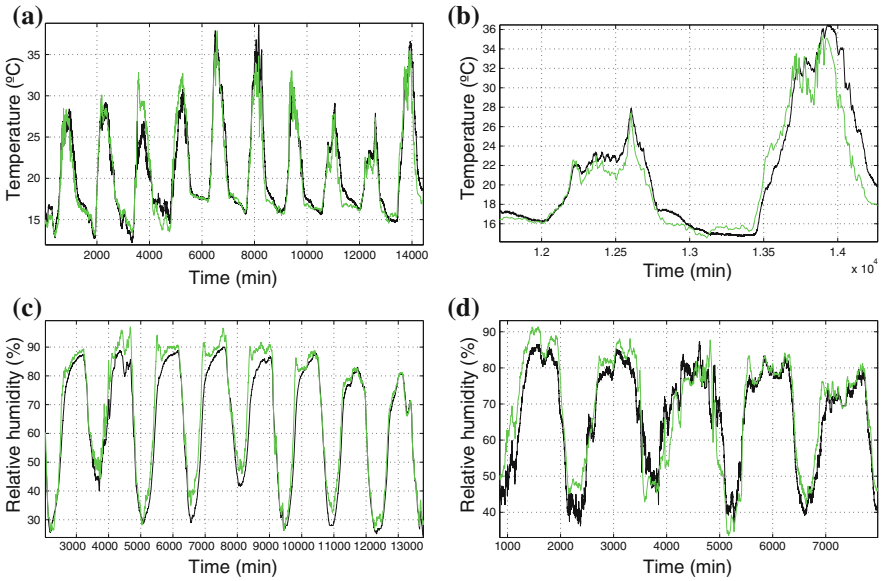

Fig. 2.12 Simulation results of Almería greenhouse with simplified physical models. **a** Temperature in April. **b** Detail of temperature. **c** Humidity in August. **d** Humidity in April

exchange of air with the outside through ventilation and the infiltration, $M_{\text{ven,a-e}}$, computed using Eq. (2.53).

2.1.2.5 Implementation, Calibration and Validation of the Model

In order to implement and calibrate de model, the same techniques described in Sects. 2.1.1.8 and 2.1.1.9 were used.

After the identification of the model parameters for the different greenhouses structures using the data corresponding to winter and summer seasons, different experiences were performed to validate this simplified climate model under different climate conditions: Winter (January), spring (April), and summer (August). Figure 2.12 and Table 2.6 show some of the results of Almería greenhouse, obtaining similar results in the other greenhouses. As it can be observed, adequate results were obtained that confirm the validity and performance of the model.

2.1.3 Data-Driven Models

As has been demonstrated in the previous sections, in the design, implementation, calibration, and validation of nonlinear simulation models, the rigorous develop-

Table 2.6 Comparative results of the estimation of the different climate variables in Almería greenhouse in the validation process with simplified physical models

	January		April		August	
	Temperature	Humidity	Temperature	Humidity	Temperature	Humidity
Variation interval	13.76–23.55 (9.79 °C)	55.64–100 (44.36 %)	14.5–38.6 (24.1 °C)	36.2–87.7 (51.5 %)	19.4–48.6 (29.2 °C)	36.36–87.4 (51.11 %)
Mean	0.74	3.64	1.31	4.32	1.14	4.71
Maximum	3.67	17.10	6.71	20.02	5.32	22.01
Standard deviation	0.71	3.23	1.36	3.86	1.09	3.79

ment of dynamic models for simulating the production system in a greenhouse is a time-consuming task that requires a wide knowledge of the involved physical processes, both in the design phase and in the model validation stage. An alternative to models based on physical principles are those obtained from data, also known as black box ones, as they are described by dynamic equations (linear or nonlinear), which coefficients are obtained through an identification procedure, defined as the problem of building mathematical models of dynamic systems based on observed data [257]. Therefore, empirical models can be developed, so that a very flexible mathematical structure with modifiable parameters estimated from experimental data can be used regardless of any consideration of the governing physical principles. The identification process begins with the design and subsequent realization of experiments in the system (using signals exciting the desired bandwidth the model has to reproduce), acquiring the necessary input and output data from the system during a given period of time. Next, the nature, size, and parametric structure of the model is determined. Based on a predetermined criteria, the model is estimated, identifying the free parameters of the selected structure. To determine whether the model is acceptable, it is then validated using real data different that those used in the parameter estimation process. It is thus an iterative process, as if the model is not properly validated, the procedure is repeated changing decisions made in the previous stages. Obviously, these models are limited as they reproduce the dynamics of a system under particular operating conditions. However, they present a number of advantages, among which the relative simplicity of obtaining the model based on an appropriate methodology stands out. In the literature, there are many techniques for obtaining data-driven models, both based on linear and nonlinear structures. In this section, some of the most used black box model structures used within the greenhouse climate framework are described.

2.1.3.1 Linear Model Obtained with Reaction Curve Method

When considering small changes around an operating point, most industrial processes can be described by a linear model, usually of high order [72]. The reason for this is that most processes are comprised of many dynamic elements, typically first order,

so that the full model order is equal to the number of elements. If, as happens in many processes, one of the time constants is much greater than the others, the smaller time constants are joined to produce a delay that acts as a pure delay. It is therefore possible to approximate the dynamic model of a complex high-order system using a first order plus dead time (FOPDT) description. Hence, one of the most common practices in process control is the estimation of simplified models of stable overdamped dynamic systems (such as the greenhouse) from type tests, being the most widespread method called the *reaction curve*, by which the system in open loop undergoes a change in the input in the form of step and so that the output is modeled as a FOPDT system, described by three parameters: Static gain c_{k}, time constant c_τ and time delay $c_{\mathrm{t_r}}$, so that the system output $Y(t)$ is described by a first-order differential equation as a function of the input $U(t)$, as described in all classical control textbooks [17]:

$$c_\tau \frac{\mathrm{d}Y(t)}{\mathrm{d}t} + Y(t) = c_k U(t - c_{\mathrm{t_r}}) \tag{2.67}$$

The main advantage of this method is its simplicity, ease of understanding by staff with little mathematical background, generally short duration of the involved test (on the order of magnitude as the dominant system dynamics), and the existence of specific control methods for this type of systems [72]. It involves introducing a step input and study the behavior of the output until steady state is reached, yielding the model parameters in a graphical manner. In [295, 407, 408, 461] climate models are obtained by this method.

This section summarizes the development of a FOPDT model of the temperature of the greenhouse (state variable $X_{\mathrm{T,a}}$, considered homogeneous) obtained using the reaction curve method. The exogenous variables and disturbances acting on the system considered for modeling purposes are the outside temperature ($D_{\mathrm{T,e}}$), wind speed ($D_{\mathrm{ws,e}}$), outside global radiation ($D_{\mathrm{rs,e}}$), and LAI (D_{LAI}). The control inputs are the vents position (U_{ven}), shade screen position (U_{shd}), and the temperature of the water inside the pipes of the heating system ($U_{\mathrm{T,heat}}$). The influence of crop on climate inside the greenhouse has been taken into account in the transfer function that relates temperature with radiation, since plants absorb part of this for their vital functions, including transpiration and thus influencing the state variable.

Analyzing the influence of each of the disturbances and control inputs on greenhouse air temperature, a series of simple models can be obtained. Different single-input single-output (SISO) models represented by transfer functions can be obtained, relating indoor air temperature with ventilation ($G_{\mathrm{ven}}(s)$), heating ($G_{\mathrm{heat}}(s)$), shade screen ($G_{\mathrm{shd}}(s)$), outside radiation ($G_{\mathrm{rs,e}}(s)$), outside temperature ($G_{\mathrm{T,e}}(s)$) and wind speed ($G_{\mathrm{ws,e}}(s)$), s being the complex variable used in Laplace transform. As transfer functions apply on linear systems, the superposition principle holds, so that the effect of each of the variables on temperature is independent and is added to produce the output (Fig. 2.13). The transfer functions are obtained by applying the Laplace transform to Eq. (2.67) with null initial conditions (defining deviation variables from a specified operating point) and have the form:

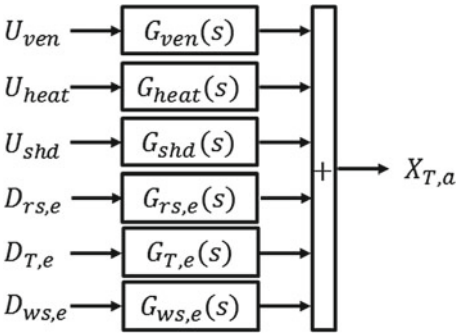

Fig. 2.13 Transfer functions relating inputs and disturbances to inside air temperature

$$G(s) = \frac{c_{\mathrm{k}}}{(c_\tau s + 1)} e^{-c_{\mathrm{tr}} s} \tag{2.68}$$

After the performed experiences, it has been observed that different parameters are obtained when steps of different magnitude or sign are applied to the inputs around a particular operating point, as was to be expected from the nonlinear nature of the system. Arithmetic means of the obtained parameters can be applied. In the Almería type greenhouse, the obtained results are summarized in Table 2.7. Notice that in the manipulated inputs, it is easy to perform open-loop step tests, but in the case of disturbances, the historical database has to be searched trying to find situations in which abrupt changes occur (with approximated step shape), while the rest of inputs and disturbances are in quasi-steady state. Thus, the obtaining of these simple models is constrained by the profile of disturbances. Obviously, if a nonlinear model has been previously developed, simple transfer functions can be obtained by linearizing it around the desired operating point. Another possibility is to obtain the parameters of the transfer functions by identifying them using, for instance, a least squares (LS) identification algorithm [257], as commented in the next section. The mean of the absolute errors obtained with these models is around 6.5 %.

Fig. 2.14 shows a graphical comparison between the real temperature for spring and winter seasons (shown in dark and continuous line) and that obtained with the simplified linear model based on the reaction curve (shown with a continuous line

Table 2.7 Parameters of the SISO FOPDT transfer functions relating air temperature with inputs and disturbances

	Radiation	Wind speed	Outdoor temperature	Ventilation	Heating	Shadow net
Static gain c_{k}	0.015 $V_{\mathrm{tsw,g}}$ (°C W^{-1} m^2)	−0.1 (°C m^{-1} s)	1 (°C °C^{-1})	−0.09 (°C $\%^{-1}$)	0.1 (°C °C^{-1})	−0.023 (°C $\%^{-1}$)
Time constant c_τ (min)	42	3	30	3	25	6
Time delay $c_{\mathrm{t_r}}$ (min)	1	1	1	1	10	1

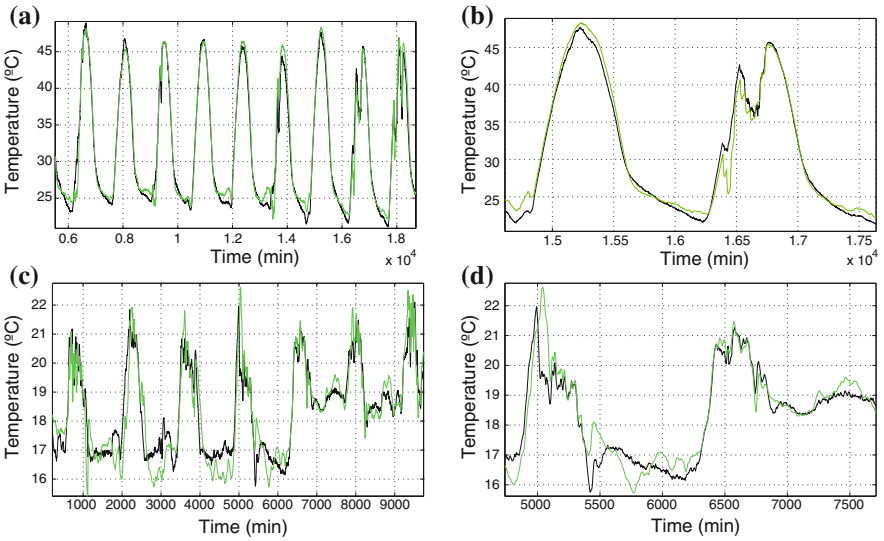

Fig. 2.14 Greenhouse air temperature simulation using FOPDT models versus real temperature. **a** Greenhouse air temperature in summer (details on the *right*). **b** Greenhouse air temperature in winter (details on the *right*)

of light color). As can be seen, the model captures the dominant dynamics of the greenhouse air temperature in different situations as clear day or the effect of the actuator continued operation. In nocturnal periods, the model shows a significant deviation from the real temperature due mainly because the effect of the thermal mass of the ground during these periods of time has not been modeled.

In order to compare the obtained results with the simulation model (Sect. 2.1.1), Fig. 2.15 shows a comparison of the real measured temperature with that simulated using linear FOPDT models and the full first principles-based model. The simplified model has worse quantitative and qualitative results, but captures the main dynamics of the system, being able to confirm the validity of the linear model obtained by the method of the reaction curve for type of applications that need some simplified models and control algorithms.

2.1.3.2 Linear Models Obtained with Input-Output Data

As shown in the previous section, due to the existence of a well-established mathematical theory and the fact that many systems present a linear behavior around certain operating points, linear models are one of the most used tools in identification for control [11]. A linear system, time invariant and causal is completely characterized by its impulse response, so that the output of the system $Y(t)$ (considering the SISO

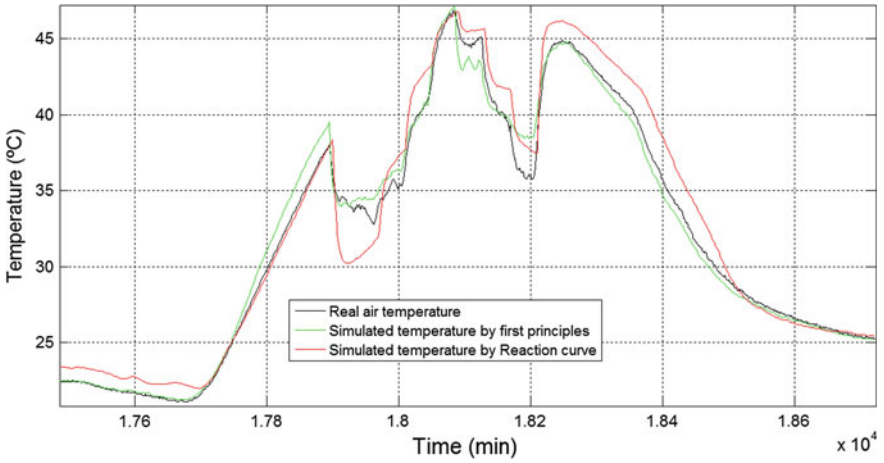

Fig. 2.15 Comparison of simulated air temperature by the complete model based on first principles versus linear model developed using the reaction curve method

case) in discrete time[1] t, is related to the input (measurable disturbance and control signal $U(t)$), through the general equation of convolution in discrete time:

$$Y(t) = \sum_{i=1}^{\infty} g(i)U(t-i) + \sum_{i=1}^{\infty} h(i)V(t-i) \tag{2.69}$$

where $V(t)$ is a zero mean white noise representing a disturbance. By applying the backward shift operator z^{-1},

$$Y(t) = G(z^{-1})U(t) + H(z^{-1})V(t) \tag{2.70}$$

where $G(z^{-1})$ is the transfer function associated to the input and $H(z^{-1})$ represents the transfer function relating the output to disturbance. Linear parametric models more widely used correspond to the following general structure:

$$A(z^{-1})Y(t) = \frac{B(z^{-1})}{F(z^{-1})}U(t) + \frac{C(z^{-1})}{D(z^{-1})}V(t) \tag{2.71}$$

For a model of this type, the transfer functions associated with inputs and disturbances are given by:

$$G(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})F(z^{-1})} \tag{2.72}$$

$$H(z^{-1}) = \frac{C(z^{-1})}{A(z^{-1})D(z^{-1})} \tag{2.73}$$

[1] In this book, t is used both for continuous and discrete time, depending on the context.

Different model structures can be obtained as a function of the polynomials which are used, see Table 2.8.

Despite the fact that the AR and ARMA models do not considered the inputs, $B = 0$, the main difference among model structures lies in the consideration of disturbances. AR and ARX structures suppose that disturbances are a white noise, $C, D = 1$, meanwhile ARMA and ARMAX suppose that they present a certain temporal structure. Furthermore, while the transfer functions of the ARX and ARMAX structures share the denominator, that is, $D, F = 1$ for both kind of model structures, OE and BJ models are totally independent. Finally, ARX, ARMAX, and BJ models are used as prediction models, while OE are simulation models which do not include any hypothesis about the structure of the disturbances since $A, C, D = 1$, and in addition, they only model the transfer function associated with the inputs of the system [257].

In [49, 51–54, 60, 104, 296, 305, 408, 461, 482, 483] climate models are obtained using these methods and different structures. In the experiences shown in this section, the system is modeled as a multiple-inputs single-output one (MISO), where the output is the inside air temperature $X_{\mathrm{T,a}}$, the disturbances are the outside temperature $D_{\mathrm{T,e}}$, wind speed $D_{\mathrm{ws,e}}$, solar radiation $D_{\mathrm{rs,e}}$, soil temperature $D_{\mathrm{T,ss}}$, and the inputs are the percentage of vents opening, one for roof ventilation $U_{\mathrm{ven,r}}$ and another one for lateral ventilation $U_{\mathrm{ven,l}}$. The models are valid in the operation around a particular operating condition, defined by the boundary conditions and state of the actuators and system. Thus, the linear models obtained by this way are only valid to operate around the particular conditions defining the data used for identification purposes.

Thus, these kind of models serve to determine seasons and stage of the crop. As an example, using data from the Almería greenhouse, the best structure fitting data using cross-validation and residual analysis was an ARX443 model of fourth order (Table 2.9), with 92.53 % fit and mean error in the order of the precision of the used temperature sensors. As an example, Fig. 2.16 shows the greenhouse air temperature predicted by this ARX model, comparing it with the real values measured in summer with a large range of variation. As can be seen, the behavior of the ARX model is reasonably closed to the data, with a mean deviation smaller than 1.5 °C. In [281, 355] there is an extensive analysis of this kind of models for both temperature and humidity identified with data obtained in different climatic conditions of the Southeast of Spain.

Table 2.8 Linear models structures

Used polynomials	Model
$B = 0, C = 1, D = 1$	Autoregressive (AR)
$B = 0, D = 1$	Autoregressive Moving Average (ARMA)
$C = 1, D = 1, F = 1$	Autoregressive with eXogeneous inputs (ARX)
$D = 1, F = 1$	Autoregressive Moving Average with eXogeneous inputs (ARMAX)
$A = 1, C = 1, D = 1$	Output Error (OE)
$A = 1$	Box-Jenkins (BJ)

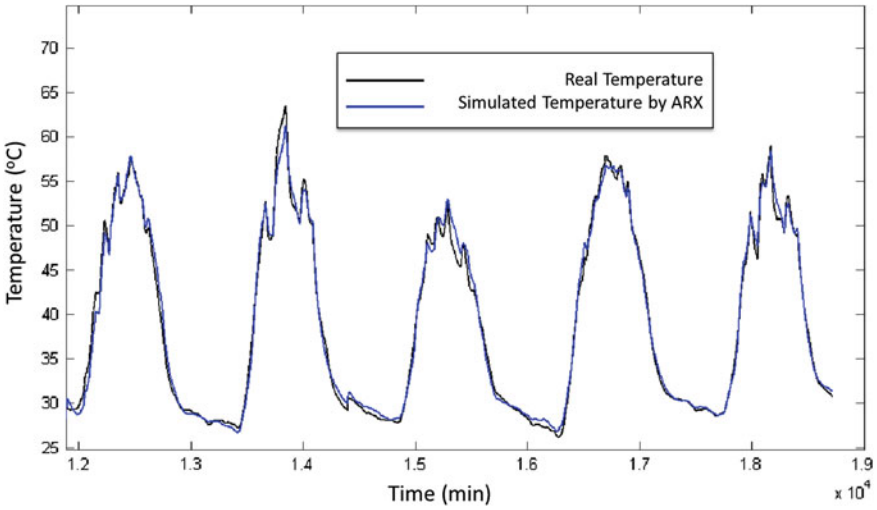

Fig. 2.16 Validation of the ARX443 temperature model with data measured in the greenhouse

Main error sources inside a wrong system identification are known as constant systematic error (bias) and random error (variance) [257]. On the one hand, constant systematic errors can be originated by: (i) input signals without adequate frequency content, (ii) a wrong choice of the model structure or operation model, for example, trying to perform the system identification using a closed loop configuration instead of an open loop one. On the other hand, random errors are introduced through the

Table 2.9 Description of the ARX model

Number of outputs: 1 ($Y = X_{\mathrm{T,a}}$)
Number of inputs: 6 ($U_1 = U_{\mathrm{ven,r}}$, $U_2 = U_{\mathrm{ven,l}}$, $U_3 = D_{\mathrm{T,e}}$, $U_4 = D_{\mathrm{ws,e}}$, $U_5 = D_{\mathrm{T,ss}}$, $U_6 = D_{\mathrm{rs,e}}$)
Order: 4 Fit: 91.53 %
A polynomial order $n_{\mathrm{a}} = [4]$
B polynomials orders $n_{\mathrm{b}} = [444444]$
delays [sample times] $n_{\mathrm{tr}} = [330000]$
Model structure: $A(z^{-1})Y(t) = \sum_{i=1}^{6} B_i(z^{-1})U_i(t) + E(t)$
$A(z^{-1}) = 1 - 1.145z^{-1} - 0.1101z^{-2} + 0.07028z^{-3} + 0.19111z^{-4}$
$B_1(z^{-1}) = -9.354e^{-5}z^{-3} - 3.428e^{-5}z^{-4} + 2.642e^{-56}z^{-5} - 4.608e^{-5}z^{-6}$
$B_2(z^{-1}) = -0.0001363z^{-3} - 7.982e^{-5}z^{-4} - 6.114e^{-5}z^{-5} - 2.5e^{-5}z^{-6}$
$B_3(z^{-1}) = 0.05246 - 0.06033z^{-1} + 0.00823z^{-2} + 0.003358z^{-3}$
$B_4(z^{-1}) = -0.0001545 - 0.01797z^{-1} + 0.00133z^{-2} + 0.0134z^{-3}$
$B_5(z^{-1}) = 5.191e^{-5} + 3.447e^{-5}z^{-1} + .857e^{-6}z^{-2} - 1.977e^{-5}z^{-3}$
$B_6(z^{-1}) = 0.6663 - 0.4621z^{-1} - 0.2932z^{-2} + 0.0933z^{-3}$
Maximum absolute error = 3.2 °C Mean absolute error = 0.7 °C Standard deviation = 0.5 °C

presence of noise in the data, which prevent that the model reproduce exactly the output of the system. In addition, models can also be affected by various factors just as: The number of parameters of the model, identification of experiment length, and the proportion between the noise/signal ratio [80].

Besides, the choice of an adequate set of input-output signals acceptable for the whole identification process is one of the most fragile points along the total procedure, since it permits a consistent estimation (free of constant systematic errors) of the parameters of the model. Within the framework of control theory, the reaction curve method is widespread used for obtaining models from data, as analyzed in the previous section. However, step or impulse signals are not always appropriate for a correct identification of industrial systems, since their frequency analysis only shows a low-frequency persistent excitation near to the stationary state. Hence, for a correct estimation of the parameters of a model, it is necessary to obtain the identification and validation data sets by means of an excitation signal with a wide frequency spectrum or within the range where the identification will be performed.

Furthermore, determining the model structure is another vital factor in order to obtain a system identification free of constant systematic errors. To do that, it is necessary to select a structure with an order high enough that helps to capture the real dynamics of the system but avoiding increasing the model order in excess. Information criteria such as the Akaike's Information Criterion (AIC) are used during the model selection stage to find a trade-off between performance and model order (between bias and variance) [257]. More information about the selection of input signals, identification data set and model structure can be found in [163, 164, 257, 350].

2.1.3.3 Linear Fuzzy Models

Fuzzy set theory uses linguistic concepts for representing quantitative values and can be used to describe the greenhouse climate based on the system identification approach [79, 148, 238, 258, 457]. Compared with traditional mathematical modeling, fuzzy modeling possesses some distinctive features, such as the reasoning mechanism in human understandable terms, the capacity of taking linguistic information from human experts and combining it with numerical data and the ability of approximating complex nonlinear functions with simple models. Several methods for fuzzy identification are proposed in the literature [239], many of which generate fuzzy rule relations from real input-output data. Generally, the resultant rule base of the fuzzy system contains a large set of rules and may make the interpretation of their consequences difficult. The Takagi–Sugeno (T-S) fuzzy model approach allows the nonlinear system to be represented under the form of a valid linear model on a restricted domain [435, 438]. This kind of models are described by rules representing the local relations of linear input-output relations in various operation points of a system. These local representations, called "submodels," make it possible to express in state space the dynamics of a system around particular operation points. Thus, the fuzzy formalism intervenes in the determination of the contribution of

each one of these submodels to the representation of the total system. That is, a nonlinear model can be represented by a set of linear models combined thorough fuzzy rules. Thus, each subsystem may contain information related to the nonlinear system. Consequently, a better resolution of the control problems is allowed.

Typically, the T-S fuzzy models represented in the discrete time state space are described by a set of N rules using membership functions μ_{li} and fuzzy variables $z_l(t)$ as follows:

$$Rule\ i:\ IF\ z_1(t)\ is\ \mu_{1i}\ \text{and}\ z_l(t)\ is\ \mu_{li}$$

$$THEN = \begin{cases} \mathbf{X}(t+1) = \mathbf{A}_i\mathbf{X}(t) + \mathbf{B}_i\mathbf{U}(t) + \mathbf{D}_i\mathbf{V}(t) \\ \mathbf{Y}(t) = \mathbf{C}_i\mathbf{X}(t) \end{cases} \tag{2.74}$$

where $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{D}_i$ and $\mathbf{C}_i$ are constant matrices of appropriate size, $\mathbf{X}(\cdot) \in \mathbb{R}^n$ is the state vector, $\mathbf{U}(\cdot) \in \mathbb{R}^m$ is the control signal vector, $\mathbf{V}(\cdot) \in \mathbb{R}^s$ is the disturbance vector and $\mathbf{Y}(\cdot) \in \mathbb{R}^p$ is the output vector. The overall global model can be structured as follows:

$$\begin{cases} \mathbf{X}(t+1) = \sum_{i=1}^{N} h_i(t)(\mathbf{A}_i\mathbf{X}(t) + \mathbf{B}_i\mathbf{U}(t) + \mathbf{D}_i\mathbf{V}(t)) \\ \mathbf{Y}(t) = \sum_{i=1}^{N} h_i(t)\mathbf{C}_i X(t) \end{cases} \tag{2.75}$$

where $h_i(t)$ are the so-called normalized activation function in relation with submodel ith such that:

$$h_i(t) = \frac{\Pi_{j=1}^{l}\mu_{ji}(z_j(t))}{\sum_{i=1}^{N}\Pi_{j=1}^{l}\mu_{ji}(z_j(t))}, h_i(t) \geq 0 \tag{2.76}$$

In [288, 290, 292], a climate model for greenhouses, expressed using fuzzy logic is presented. In particular, a T-S model is derived from a standard nonlinear model representing energy and water vapor balances (from the pseudo-physical climate models showed in Sect. 2.1.2), giving a set of linear models related through fuzzy logic. This makes it possible to derive a greenhouse climate model based on linear models. The variables considered for modeling purposes are:

- Output: Inside air temperature ($X_{\mathrm{T,a}}$) and humidity ($X_{\mathrm{H_{a,a}}}$).
- Input: Aperture of the roof ($U_{\mathrm{ven,r}}$) and lateral ($U_{\mathrm{ven,l}}$) ventilations and heating system ($U_{\mathrm{T,heat}}$).
- Disturbances: Outside temperature ($D_{\mathrm{T,e}}$), wind speed ($D_{\mathrm{ws,e}}$), soil surface temperature ($D_{\mathrm{T,ss}}$), outside global solar radiation ($D_{\mathrm{rs,e}}$) and LAI (D_{LAI}).

The model was validated with data of winter and summer from quantitative and qualitative viewpoints, because it follows the dynamics of the modeled variables (see [292]) and the errors are within acceptable intervals for this type of applications (the estimation root mean square error, RMSE, of the temperature is 1.24 °C whereas for humidity is 10.6 %). As an example, Fig. 2.17 shows the interior temperature predicted by the T-S model, compared with the real values measured in early spring

(the first eight days shown) and late autumn (the last four days) [291]. The T-S model is reasonably close to the data taken in the different climatic conditions, with a mean deviation smaller than 1 °C.

It can be seen that during some nights the fuzzy model can not exactly approximate the air temperature inside the greenhouse. This drawback often attributed to T-S approximation method is in reality a problem of the mathematical model which is used to extract the T-S fuzzy model.

2.1.3.4 Nonlinear Volterra Models

Before using nonlinear models, it is always advisable to explore all possibilities of simplicity. After linear models, the next step in complexity are those models with concentrated nonlinearities such as Volterra, Hammerstein or Wiener ones [114]. These models are unions of linear dynamic and static nonlinear blocks. Specifically, Volterra models were used to generically exhibit a good behavior and their structure can be exploited in the design of controllers, especially in the case of second-order models with the truncation of terms (truncation orders N_1 and N_2), which can be defined as:

$$Y(t) = h_0 + \sum_{i=1}^{N_1} a(i)U(t-i) + \sum_{i=1}^{N_2}\sum_{j=1}^{N_2} b(i,j)U(t-i)U(t-j) \qquad (2.77)$$

which corresponds to a linear convolution model with a nonlinearity as additional and additive term, as described in [156]. In that model, $Y(t)$ and $U(t)$ represent the last measured output and input to the system, respectively (t is the actual sampling instant). The offset is denoted with h_0 and the linear and nonlinear term parameters are given by $a(i)$ and $b(i, j)$, respectively. Notice that Volterra models are frequently used to model bilinear systems in such a way that it seems to be a good idea to use this formulation for modeling greenhouse temperature dynamics, including the disturbances in the nominal formulation of second-order Volterra models.

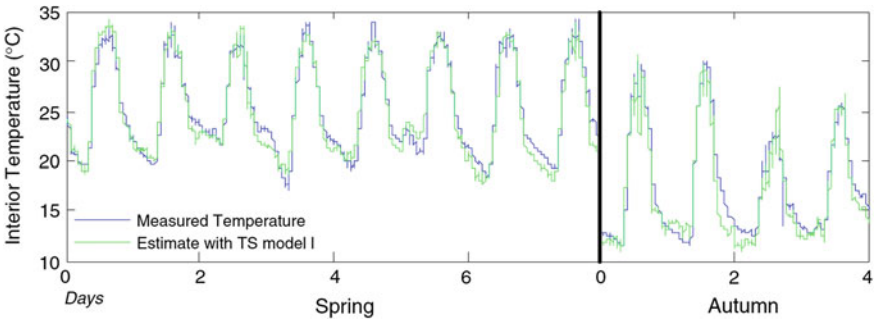

Fig. 2.17 Validation of the TS model with data measured in the greenhouse. As a courtesy of the authors [291]

A preliminary Volterra model was developed to model the inside temperature of an empty greenhouse (without crop) in order to evaluate the behavior of this modeling technique for this kind of systems [155]. Several pseudo random multilevel sequence (PRMS) [350] tests were performed using natural ventilation to obtain adequate data for identification purposes, because typical PRBS (pseudo random binary sequence) and RBS (random binary sequence) tests do not sufficiently excite nonlinear systems. The resulting model adequately fitted the real data but the number of parameters was excessively high. Furthermore, the crop has an important effect on the greenhouse temperature and thus it is a key factor to be included in the system model. In [156], two Volterra models (AR and Non-AR) are developed in order to account for the main dynamics describing changes in inside air temperature to outside weather using only natural ventilation. The influence of the crop is taken into account as a disturbance to the greenhouse temperature by means of the LAI. Thus, the main variables considered for modeling purposes are:

- Output: Inside air temperature ($X_{\mathrm{T,a}}$).
- Input: Aperture of the roof ($U_{\mathrm{ven,r}}$) and lateral ($U_{\mathrm{ven,l}}$) ventilations.
- Disturbances: Outside temperature ($D_{\mathrm{T,e}}$), wind speed ($D_{\mathrm{ws,e}}$), soil surface temperature ($D_{\mathrm{T,ss}}$), outside global solar radiation ($D_{\mathrm{rs,e}}$) and LAI (D_{LAI}).

The main interest is to see how these models cope with the nonlinear behavior inherent in the relationship between temperature and vents aperture, through the ventilation rate, which is one of the most difficult dynamics to be modeled in the greenhouse.

A second-order Volterra series model of the greenhouse temperature was identified. In the model validation, carried out with the second data set containing the period from September 2008 to June 2009, a mean square error in the temperature of 0.93 was obtained. As a representative result, a comparison of the greenhouse temperature and the output of the identified model is given in Figs. 2.18 and 2.19, both for the autumn (10–19 January 2009) and spring conditions (15–24 May 2009), respectively. As can be seen in the results, the model output shows a promising fit with the measured greenhouse temperature. In autumn conditions, the model presents a mean value of the absolute error of 0.6 °C, a standard deviation of 0.5 °C, a mean relative error less than 4 %, and a maximum error of 2.1 °C in a range of 11.1–26.4 °C. In spring conditions, similar results were obtained, resulting in a mean value of the absolute error of 0.68 °C, a standard deviation of 0:63 °C, a mean relative error less than 4 %, and a maximum error of 2.5 °C in a range of 15.4–31.4 °C.

2.1.3.5 Nonlinear Neural Networks Models

Artificial neural networks (ANN) are computational elements inspired by networks of neurons in the nervous system of living beings. They consist of elements (neurons or nodes) connected in parallel, whose collective action is able to reproduce complex functions. In addition, the connections between nodes are customizable so that the overall function of the network can be modified [11]. These characteristics can be used to solve problems of identification of dynamical systems like climate

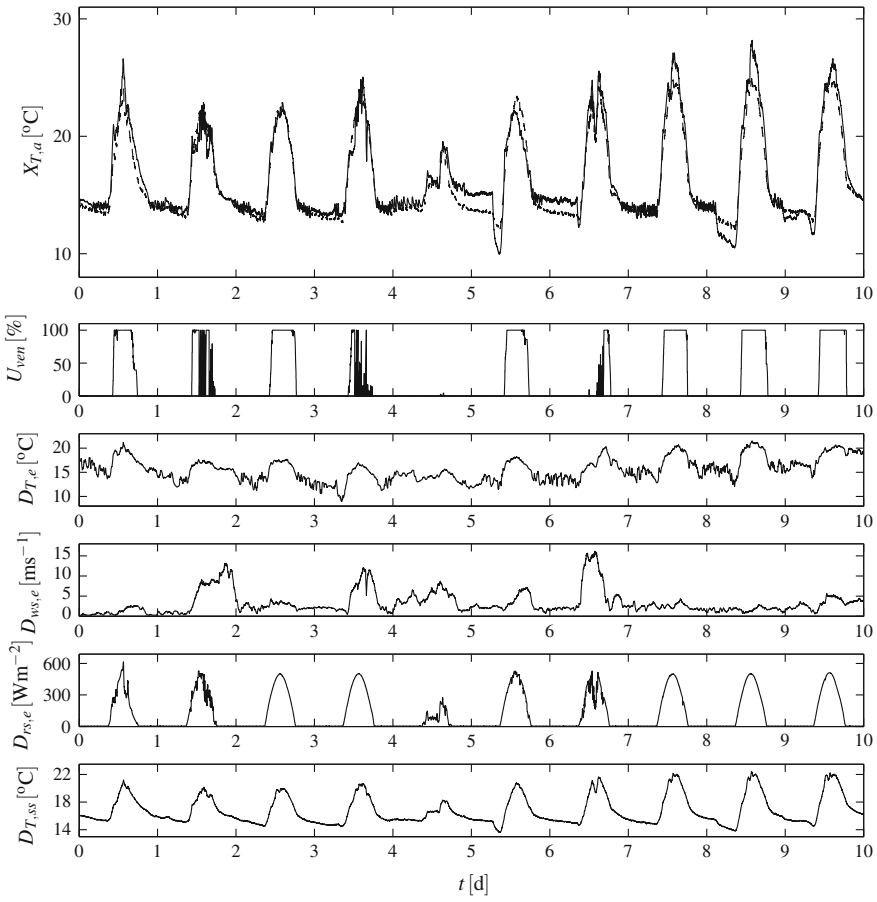

Fig. 2.18 Data set used for the model validation with the greenhouse temperature $X_{\mathrm{T,a}}$ (*solid line*) and the model output (*dashed line*) for autumn conditions (10–19 January 2009), the input U_{ven} (aperture of the roof and lateral windows) and the disturbances $D_{\mathrm{T,e}}$ (outside temperature), $D_{\mathrm{ws,e}}$ (outside wind speed), $D_{\mathrm{rs,e}}$ (outside global solar radiation), and $D_{\mathrm{T,ss}}$ (soil surface temperature). As a courtesy of the authors [156]

that is generated inside a greenhouse, finding numerous references in the literature emphasizing specialized work [8, 128, 129, 131, 254, 255, 312, 380, 381, 440, 453, 471].

The reasons for using neural networks as identification systems is due to its fast response (parallel processing), their ability to interpolate, their flexibility to describe nonlinear functions and the ability to work with spaces of large dimension. Different neural models for the identification of dynamic systems exist, most of them are extensions to the nonlinear case of linear parametric models, although the NARX model (nonlinear autoregressive model with exogenous inputs) provides great flexibility and allows it to be adjusted with simple algorithms [11]. The output of the

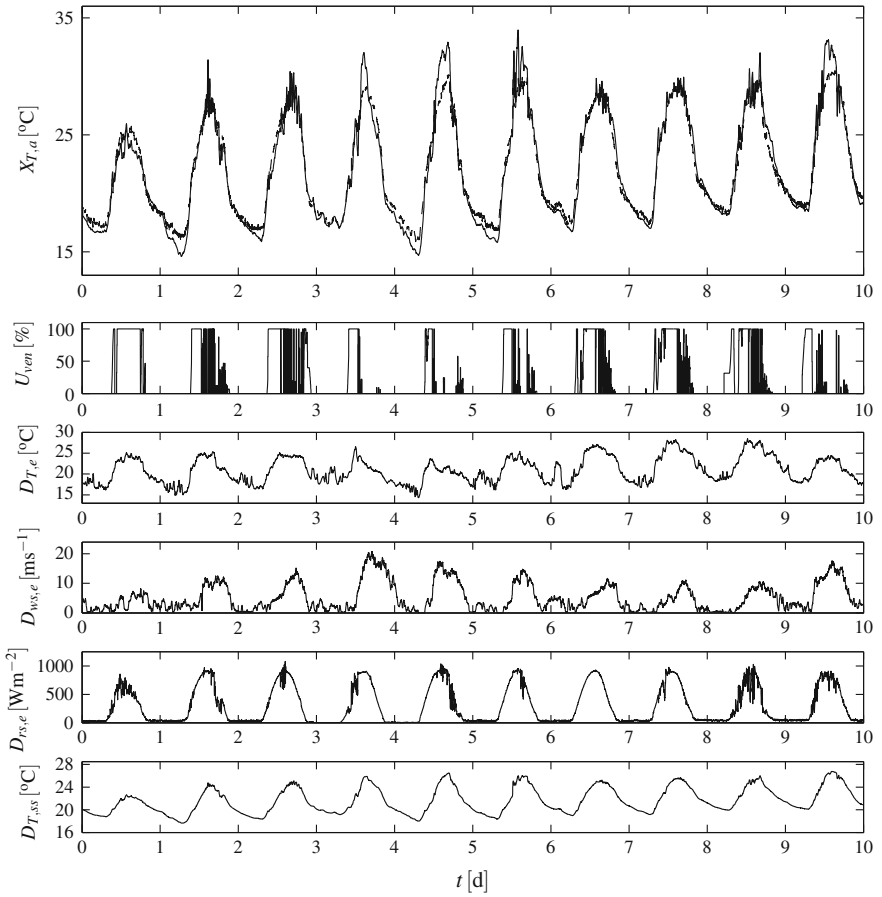

Fig. 2.19 Data set used for the model validation with the greenhouse temperature $X_{\mathrm{T,a}}$ (*solid line*) and the model output (*dashed line*) for spring conditions (15–24 May 2009), the input U_{ven} (aperture of the roof and lateral windows) and the disturbances $D_{\mathrm{T,e}}$ (outside temperature), $D_{\mathrm{ws,e}}$ (outside wind speed), $D_{\mathrm{rs,e}}$ (outside global solar radiation), and $D_{\mathrm{T,ss}}$ (soil surface temperature). As a courtesy of the authors [156]

NARX model requires past values of input and outputs, so that tapped delay lines (TDL) are used in the implementation. Mathematically the prediction model is given by the following expression [356]:

$$\mathbf{Y}(t+1) = h[\mathbf{Y}(t), \ldots, \mathbf{Y}(t-l), \mathbf{U}(t), \ldots, \mathbf{U}(t-m), \mathbf{D}(t), \ldots, \mathbf{D}_m(t-n)] \quad (2.78)$$

where $\mathbf{U}(t)$ is the input vector of the system at discrete time instant t, which includes the values of ventilation, heating and shade screen, $\mathbf{Y}(t)$ is the output vector at time instant t, which includes the values of temperature and relative humidity, $\mathbf{D}_m(t)$ is the vector of measurable disturbances at time t, which includes the values of outside

Table 2.10 Past values of the inputs of the ANN

Variable	Past values
Temperature	3
Humidity	3
Outside temperature	4
Solar radiation	2
Wind speed	9
Ventilation	5
Heating	10
Shade screens	2

temperature, humidity, radiation, wind speed and direction. The orders l, m and n of the output, input and disturbance vectors are known only in some situations, but generally are obtained by observations made on the system. The NARX model shown in Eq. (2.78) can be implemented by neural networks using as input vector of a historic values of the measured variables [356].

The time domain has been considered trying the past values of the variables as different inputs to the system that feed a static neural network by means of TDL. The number of past values used, as a rule, is unknown and difficult to determine. For a higher number of past values, the stronger the prediction, but at the same time, the model may be inefficient due to the high number of entries required. Moreover, a lower number will cause the model could not accurately predict future outputs. There are a number of methods for selection of input variables for nonlinear models, having used a model based on the estimated gradient as a ratio of distances between points in the enclosed space entries, combined with the use of the information obtained from linear methods using ARX type models, so that the past values of input and output feedback linear models are used as indicative values for the neuronal nonlinear models [356]. For example, the last values used to obtain a model of the temperature and humidity inside the greenhouse Araba number 2 in a given instant t, are shown in Table 2.10 [355, 356]. The selected ANN structure is a multilayer perceptron (MLP), with an input layer with 38 nodes, one hidden layer with 8 nodes and nonlinear activation, and an output layer with 2 nodes. The weights are set as connections to nodes with constant values. The number of nodes is determined by training the different networks and determining, by the method of cross-validation, their approximation and generalization ability. This process was carried out using two disjoint sets of data: One for training the network and one for validation. The training of the implemented neural network has been performed using the method of supervised learning with back propagation, so that the least squares criterion is minimized, as a function of the square of the difference between the measures of real variables acquired in greenhouses and the values estimated by the model, as has been done in the calibration of other models developed in previous sections.

In order to validate the ANN model of temperature and relative humidity using MLP as a network representation, data from Araba number 2 greenhouse and from

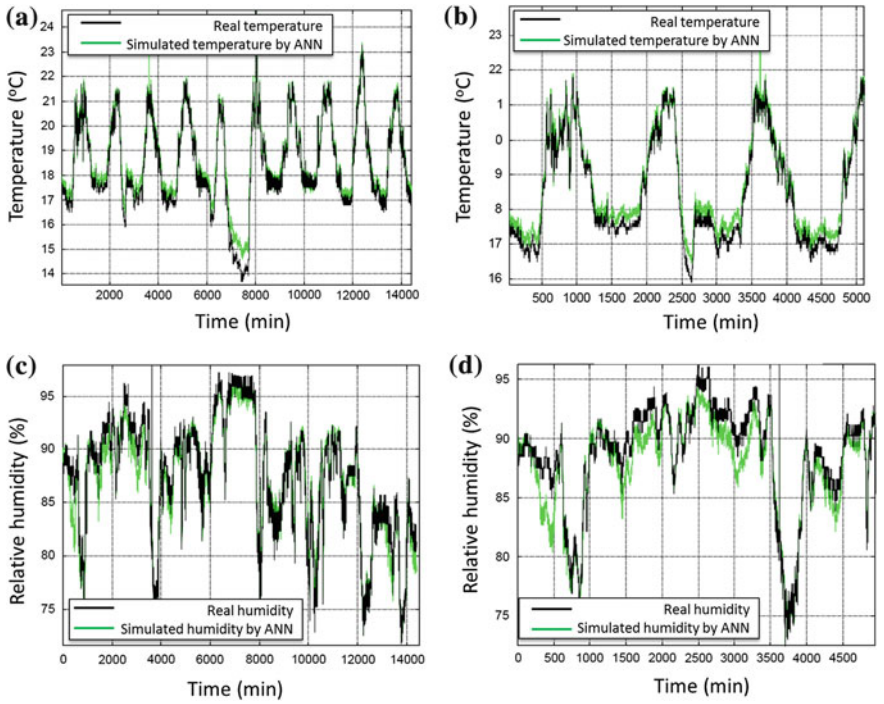

Fig. 2.20 Comparison between real temperature and ANN model estimation. **a** Temperature evolution during 10 days. **b** Zoom of the evolution. **c** Relative humidity evolution during 10 days. **d** Zoom of the evolution

the same season in which the model is calibrated (winter) were used. The results obtained for the temperature are shown in Fig. 2.20 in which the real measurements are represented by a dark solid line and the estimated by the models by a light solid line color.

As can be seen, the temperature in summer is overestimated mainly because the greenhouse roof is whitened which affects the final result. Table 2.11 shows the statistical of the residues in absolute value of the real temperature and relative humidity and the estimated by the neural model. The temperature relative error is 2.4 % in winter, while rises above 6 % for the summer. It is due to the fact that the neural network has been trained with data of winter, so the best results are obtained with conditions an data located near the area of the input/output data used in the identification processed. The extrapolation to distant points from the training space may not produce the desired results [11], as seen in the results for summer whose outside disturbance are different to the training space formed by the corresponding disturbances to winter. However, the relative humidity model behaves correctly in all tests performed at different season of the year, with a constant error for all of them, not being more than 5 % in the range of variation of this variable. Even so, the results are acceptable for use in model-based control algorithms.

Table 2.11 Results with the ANN model

	Temperature			Relative humidity		
	January	August	August	January	August	August
Variation interval	13.76–.55 [9.79 °C]	19.4–48.6 [29.2 °C]	21.1–49 [27.9 °C]	55.64–100 [44.36 %]	36.36–87–47 [51.11 %]	21–94 [73 %]
Mean error	0.24	1.75	1.62	1.96	2.46	3.24
Maximum error	2.59	3.72	3.27	1.12	2.09	2.51
Standard deviation	0.21	1.12	1.02	1.12	2.09	2.51

Another possibility is to use finite impulse response (FIR) discrete time nonlinear models with integrated variables for greenhouse indoor temperature simulation in order to reduce the number of past values needed as inputs [12]. In this case, the interest is in obtaining a discrete time model for simulation where the actual output of the system $Y(t)$ is not known at any time (except perhaps at the beginning of the simulation $t = 0$). A model that uses past values of Y (i.e., an autoregressive model) must then use its own output $\hat{Y}$ as an estimation of the true output and use it recursively during the entire simulation time. This can cause a built-up of the simulation error producing errors that are larger as the simulation horizon increases. A model that uses just past values of the input signal belongs to the family of FIR models. The output of a FIR model is obtained as a linear combination of past values of the system's input. Since the real output of the system is not needed, this kind of models produces simulation errors that are independent of the simulation horizon. Also, any FIR model obtained by identification is stable, in the bounded input—bounded output (BIBO) sense, since the output of the model is obtained as a combination of past input values. Due to the nonlinear behavior of the greenhouse temperature, NFIR (nonlinear counterpart of the FIR family) was used. The input variables used are the same, but in the first case, the function that combines them is a nonlinear mapping produced in this section by an artificial neural network. As a result, the input vector for the ANN at sample time t is computed as:

$$\mathbf{U}(t) = [U_1(t - d_1 - 1), \ldots, U_1(t - d_1 - n_1), \ldots, U_{nv}(t - d_{nv} - n_{nv})] \quad (2.79)$$

where all variables included (U_1 to U_{nv}) are control actions or disturbances. The values d_j $j = 1, \ldots, nv$ are the dead time for variable j. Model orders are the number of lagged values n_j of each variable from $j = 1$ to $j = nv$. The dimension of the input vector is thus $\dim \mathbf{U} = \sum_{j=1}^{nv} n_j$.

The output variable is assumed to be a nonlinear function of the input vector plus a white noise signal n.

$$\mathbf{Y}(t) = f(\mathbf{U}(t)) + n(t) \quad (2.80)$$

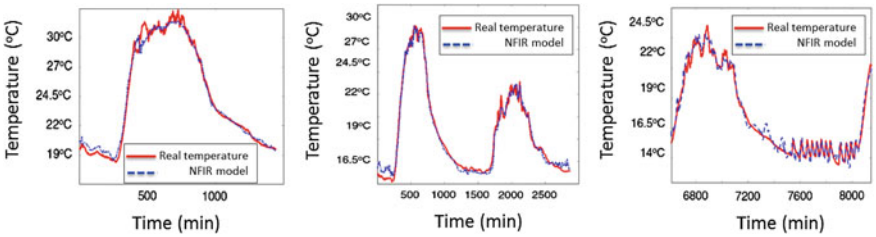

Fig. 2.21 Comparison between real temperature and ANN model estimation. As a courtesy of the authors [12]

A variation of NFIR models consists of including integrated values of some variables. Then, instead of using $U(t-d-i)$ for $i=1,2,\ldots,n$ it is possible to accumulate the effect in just m sums with $m<n$ being the new variables

$$w(t,j)=\sum_{i=s_j}^{i=f_j} U(t-d-i) \tag{2.81}$$

for $j=1$ to $j=m$. Obviously the initial and final index for the sums must verify: $s_1=1, s_{p+1}=f_p+1$ for all $p=1,\ldots,m-1$ and $f_m=n$. The initial (s) and final (f) index define a time window in which the integration of variable U takes place to yield the integrated variable W. Using the integrated variables, the input vector for a neural NFIR model at sample time t is computed as:

$$\mathbf{U}(t)=[w_1(t,1),\ldots,w_1(t,n_1),\ldots,w_5(t,1),\ldots,w_5(t,n_5)] \tag{2.82}$$

The adjustable parameters of the models are obtained minimizing a quadratic criterion of the simulation error. For NFIR12314 model structure (the root mean squared error was the lowest), a neural network was constructed of 20 hidden nodes, obtained better results than the complete first principles model [12]. Figure 2.21 shows a simulation example. The real temperature is plot in solid line, while the simulated one obtained by model NFIR12314 is plot in dashed line. The simulation corresponds to: a. clear day, b. vents opening, c. heating during the night.

2.2 Crop Growth Models

2.2.1 Tomato Growth and Development Models

Growth can be defined as an increment in biomass or an increment in the dimensions of the plant, that are quantitative aspects [87]; growth can also be defined as an increment in weight or height of the organs of the plant [149]. Development is a

concept that indicates a change or organized process (not always) towards a superior state, more organized or more complex [45]. Development implies qualitative aspects that are not only phase transitions as the change from juvenile to adult state, but also the formation of new organs, the senescence of the organs [87] or the start of the establishment of fruits (fruit setting), or tuber filling, changes in appearance of the plant that can occur even if there is no weight increment [149]. Development is an irreversible process of change in the state of an organism, which generally progresses according to a pattern more or less fixed and specific for the species [149]. In [272] there is a very detailed numeric system which allows the description of the states of phenologic development through a uniform code in plant species; this system is known as the BBCH scale (Biologische Bundesanstalt, Bundessortenamt and Chemical).

To study growth and development of the crops, it is necessary to understand the physiological processes behind them, as far as the current knowledge allows it. The basic physiological processes in plants are: Photosynthesis, respiration, metabolic activities, nutrients and water uptake, nutrients and water transport, transpiration and the generation of reproductive structures. Many of these processes have their limits genetically determined, but microclimate, substrate, nutrition regime, and some specific enzymes play an important role [276].

Because of its importance, microclimate has been extensively studied in the modeling of growth and development of greenhouse crops. Microclimate includes main elements that affect the physiological processes of the plant: Solar radiation, CO_2 concentration and temperature. The considered radiation is the one that has the range utilized by plants which is the PAR. There are other limiting elements that play a role as relative humidity. Some authors give different weight to the elements mentioned before, for example, for Challa and coworkers [87], the most important climatic factors within the greenhouse are: CO_2 concentration, air temperature, and vapor pressure of water. Radiation can be considered as a surrounding condition due to the fact that it is imposed by the exterior climate.

2.2.1.1 Importance and Classification of Growth Models

Models as an abstraction of reality are a tool that humans have developed in many disciplines and also, with some delay, in the food production field. Is in the industry where models have had a huge development, specially compared with agriculture. At the beginning of the 1970s of the last century, the perception of the development of models in agriculture can be summarized as follows: “A chemical engineer would not design a chemical plant, nor its control processes, without first having a model of the chemical process to be done by such plant as the foundation for the design. However, the agricultural engineer, who is in charge of the design of environmental systems for the biggest chemical factory in the world (the transformation of the energy from the light and other chemical processes to food), does not have a proper model of the system with which he or she works” [101].

This situation has changed as it is explained below; however, agriculture is still a field where the time between the generation and the application of a new technology is greater than that in the services industry [338].

In agriculture, there are several families of models: Descriptive, teleonomic, process based and functional-structural. The descriptive models include the statistical regression and the empirical or black box ones. As has been treated in previous sections, they are direct descriptions of data and they indicate the relationship between variables of a system, but do not give any explanation about mechanisms for those relationships.

The models based on processes, also called explanatory or mechanistic, contain submodels with at least one hierarchical level of greater depth than the described response [243]. In a physiological model, every additional depth level increases the explanatory power of the model. The mechanistic modeling follows the reductionist traditional method, which has been successfully applied in the Physical Sciences, Molecular Biology and Biochemistry [451].

The empirical models are direct descriptions of observed data, which can be of great utility in certain circumstances [451]. In an empirical model, any proposed mathematical relationship is not restricted by physical laws such as the conservation of energy or the laws of thermodynamic or by biological information, or by any knowledge of the structure of the system [451].

Another approach to modeling are the teleonomic models, which are clearly formulated in terms of goals [451]. Even though this view has been questioned, some authors claim the importance of these models to model the processes in live organisms, and they indicate that processes oriented with an objective are intrinsic to life itself, and not to nonliving things [332]; therefore, these types of models can be useful as a link between empirical models and explanatory models [451], and they have been applied in many aspects, among them the distribution of dry matter between root and shoot [470] and to cellular level modeling [208].

Another approach are the functional-structural models, these models are oriented to merge geometrical models of plant visualization with process based models. In this approach, the goal is to control the whole plant development in its organogenesis and photosynthesis; the organs play the true roles as sources and sinks and have interaction between the architecture and the functioning during the plant development [344], this approach has emerged relatively recently and represents one of the key challenges for plant modeling [445].

Most of the explanatory models are based on photosynthesis. The main components of the models based on photosynthesis are: Development of leaf area, light interception, photosynthesis, and respiration [265].

Models in crops have several applications. It is possible to utilize them in help systems for decision-making in agricultural production, in scientific research, in the definition of politics for agricultural development, in agricultural teaching [137], and also in the climatic control of greenhouses [360, 430, 441].

2.2.1.2 Growth Models for Tomato in Greenhouses

In the tomato crop, the most important growth and development models are of the explanatory type, they are based on physiological processes; besides, these models have been validated in different degrees and varied conditions of the crop. Despite the improvement of this type of models, there is still a lot to be done, and the most important weaknesses of the explanatory models are: The simulation of the development of leaf area, the maintenance respiration, organ abortion, the content of dry matter and the quality of the product [185, 265]. In [185], it is pointed out that quality modeling in dry matter is a very important parameter. The models are described in the next paragraphs.

Tomgro is a physiological model of development and yield for the tomato crop, in which a series of differential equations represents the changes in number and weight of leaves, fruits, segments of stem, leaf area as well as initiation of new organs, their age, senescence, or those that are pruned. The model utilizes an approximation source–sink for the distribution of carbohydrates for the growth of different organs [211].

This model is schematic and it is also modular, which means that it can be easily adapted and its subprograms can be replaced by others and it can be combined with more understandable greenhouse models, and it can also be utilized in procedures of economical optimization [105]. This model was calibrated and validated with data acquired in controlled conditions for varieties of "indeterminate" growth type [106, 211].

Tomgro has been modified to include the simulation of the growth and development of individual organs, providing good simulations of number and weight of fruits per cluster [211]. This model has also interface adjustments that permit the establishment of initial parameters and conditions before the simulation [136]. In the most complete version, *Tomgro* can have 574 state variables and simulates with great detail the development of fruits due to the fact that every fruit has a specific position within the cluster, and in relation to the number of clusters [216]. With the aim of adapting *Tomgro* for the climatic control of greenhouses, the model has been reduced to five state variables, trying to preserve its main elements that allow it to be an explanatory model [212].

De Koning [227] developed a model to predict the distribution of dry matter in tomato, which can have 300 state variables. The number of organs is evaluated through the prediction of initiation, abortion and harvest of individual organs. The model calculates the sink strength of each organ through the potential growth. It is capable of predicting in a reasonable way the formation of clusters, time frame for the growth of the fruit and the distribution of dry matter, although the prediction of number of fruits per cluster does not give very acceptable results.

Tomsim is another model developed for tomato of the explanatory type with modular structure, which simulates growth and development [181–184]. The production of dry matter in this system is predicted by a general growth model for greenhouse crops, which has as a foundation the estimation of photosynthesis proposed by Gijzen [146]

that was validated by Heuvelink [184]. The functions of fruit development were adapted from the model developed by de Koning [227].

The *Tomsim* model of production of dry matter and its distribution to leaves, stems, and fruits was validated with different transplant dates and plant density, and the model was completed with data sets from commercial greenhouses, which is important to the fact that the tomato crop covers a complete season, whereas the experiments are evaluated only until 100 days after transplant. Besides, the model helps to perform analyses for the tomato crop and it can be a significant contribution as a decision support system in the crops management [182].

The *Tompousse* model is aimed to simulate the weekly production of greenhouse tomato taking into consideration the information available according to the production conditions. The key stages for the making of yield are the average transmission of radiation for the cover of the greenhouse, the interception of radiation by the canopy (dependant on LAI), the conversion of radiation to dry matter (in particular dependant on the amount of CO_2 and also on the distribution of a fraction of dry matter to the fruits). The model allows the user a good simulation of the production curves in changing climates as that in the French Brittany and in the Mediterranean region [135].

In [468], a simulation model for tomato crop was calibrated and validated using data from Spain and Netherlands in order to use this model in a model-based method to design greenhouses.

In what follows, the units of variables are not included in the paragraphs for the sake of space, but can be found in the acronyms section and in the tables included in the following sections.

2.2.1.3 The Simplified *Tomgro* Model

The simplified *Tomgro* model emerges as an option to eliminate the complexity of the complete *Tomgro* model [211] and to give the possibility of using the model in control systems on line preserving its physiologic characteristics [212]. The state variables in this model are: Number of nodes (X_{N}), leaf area index (X_{LAI}), total dry weight (X_{W}), dry matter of fruits (X_{F}), and dry mater of mature fruits (X_{MF}). For more details about the simplified *Tomgro* model refer to [212].

The number of nodes is the result of the speed of nodes formation and this is a function in sections of the variable temperature of the greenhouse microclimate, modulated by an empirical coefficient. The maximum speed in the nodes appearance is established under temperature conditions between 12 and 28 °C and it is considered that a temperature lower than 9 °C or greater than 50 °C stops the nodes appearance. This state variable is calculated in the same way in all *Tomgro* versions [212].

The LAI considers daily average temperature, empirical coefficients, and plant density. When all the leaves within the plant reach their maximum, they will be pruned or will enter the state of senescence which is also considered in the model.

The total dry matter is a function of the growth rate of the plant. This growth is a function of photosynthesis rate minus respiration multiplied by a conversion

coefficient from carbohydrates to dry matter and multiplied also by a function of distribution of dry matter to roots, which depends on the number of nodes. When the maximum LAI is reached, a coefficient of lost of dry matter is applied. Photosynthesis is calculated with the variables: Temperature, photosynthetically active radiation, carbon dioxide, and LAI. Respiration is a function of temperature and total dry matter.

Dry matter of fruits starts from the amount of nodes in which appears the first fruit. Parameters of allocation to fruits and transition from vegetative to reproductive growth are included. It is also included a function that calculates the effect of the daily average temperature on the distribution between vegetative and reproductive growth. Finally, it is considered a critical temperature for warm days, above such temperature the allocation to fruits decreases.

Regarding dry matter of mature fruits, the dynamic of mature fruits is based on the effect of temperature over fruit ripening through a linear function in sections that is activated at certain amount of nodes, that indicates the period from the appearance to the ripening of the first fruit. The assumption in the model is that mature fruits are harvested immediately.

With data from greenhouses located Southeast of Spain, a region with mild climate conditions with a minimum temperature of 12 °C, without addition of carbon dioxide and for a fall-winter season, a process of parameter adjustment was performed using the LS method with data sampled every minute.

The main equations of the *Tomgro* model [211, 212, 341] (some parameters that have to be calibrated with their units are explained in Table 2.12) are given by:

$$\frac{\mathrm{d}X_{\mathrm{N}}}{\mathrm{d}t} = N_m f_N(X_{\mathrm{T,a}}) \tag{2.83}$$

where $f_N(X_{\mathrm{T,a}})$ is a piecewise linear function that depends on temperature [212]. The dynamic evolution of LAI is given by:

$$\frac{\mathrm{d}X_{\mathrm{LAI}}}{\mathrm{d}t} = \begin{cases} \rho \delta_l f_{\mathrm{LAI}}(\bar{X}_{\mathrm{Td,a}}) \frac{\exp(\beta_l(X_N - N_b))}{1+\exp(\beta_l(X_N - N_b))} \frac{\mathrm{d}X_N}{\mathrm{d}t} & \text{if } X_{\mathrm{LAI}} < c_{\mathrm{LAI,max}} \\ 0 & \text{if } X_{\mathrm{LAI}} \geq c_{\mathrm{LAI,max}} \end{cases} \tag{2.84}$$

where ρ is the plants density, and $c_{\mathrm{LAI,max}}$ is defined as the LAI when the set of leaves of the plant reaches its maximum (it will be pruned or will inter into state of senescence) and unitless function $f_{\mathrm{LAI}}(\bar{X}_{\mathrm{Td,a}})$ depending on average daily temperature reduces the rate of leaf area expansion.

The total dry weight is described by:

$$\frac{\mathrm{d}X_W}{\mathrm{d}t} = GR_n - p_1 \rho \left(\frac{\mathrm{d}X_N}{\mathrm{d}t} \right) \tag{2.85}$$

where p_1 is a parameter describing loss of leaf dry weight per node after reaching $c_{\mathrm{LAI,max}}$ and GR_n is a function modeling net aboveground growth rate, defined as

$$GR_n = c_E(V_{\mathrm{fot}} - V_{\mathrm{res}})(1 - f_R(X_N)) \tag{2.86}$$

Table 2.12 Estimated parameters for *Tomgro* reduced model

Parameter	Description	Value	Units	Variable
c_E	Coefficient of dry matter conversion	0.12	$(g_{\text{dry weight}}\ g^{-1}_{CH_2O})$	X_W
c_{extlw}	Light extinction coefficient	0.61	–	X_W
N_m	Maximum rate of nodes appearance	0.57	(node d^{-1})	X_N
N_b	Parameter in expolinear equation	7	(node)	X_{LAI}
T_{crit}	Mean daytime temperature above which fruit abortion starts	26	(°C)	X_F
$V_{\max}$	Maximum increase in vegetative tissue per node	6	$(g_{\text{dry weight}}\ \text{node}^{-1})$	X_{LAI}
α_F	Maximum partitioning of new growth to fruit	95	(fraction d^{-1})	X_F
α_l	Light efficiency	0.09	$\left(\mu\text{mol}_{CO_2}\ \mu\text{mol}^{-1}_{\text{absorbed photon}}\right)$	X_W
β_l	Coefficient in expolinear equation	0.5	(node^{-1})	X_{LAI}
δ_l	Maximum leaf area expansion per node	0.062	$\left(m^2\ \text{node}^{-1}_{\text{leaf}}\right)$	X_{LAI}
ν	Vegetative-fruit transition coefficient	0.38		X_F
τ_c	Carbon dioxide efficiency	0.12	$(g_{\text{dry weight}}\ \text{node}^{-1})$	X_W

where V_{fot} is photosynthesis, V_{res} is respiration, c_E is the growth efficiency, a parameter that expresses the conversion of carbohydrates to dry matter and $f_R(X_N)$ is the fraction of distributed growth to roots and it is considered as a function of the number of nodes.

Photosynthesis is given by the following equation:

$$V_{\text{phot}} = \frac{c_{\text{cnv,phot}} F_{\max} f_{\text{T,phot}}(X_{\text{T,a}})}{c_{\text{extlw}}} \ln\left[\frac{(1-c_m)F_{\max} + \alpha_e c_{\text{extlw}} V_{\text{PAR}}}{(1-c_m)F_{\max} + \alpha_e c_{\text{extlw}} V_{\text{PAR}} \exp\left(-c_{\text{extlw}} X_{\text{LAI}}\right)}\right] \tag{2.87}$$

where c_m is the light transmission coefficient through leaves, c_{extlw} is the light extinction coefficient, α_e is the light efficiency, $c_{\text{cnv,phot}}$ is a units conversion coefficient, $f_{\text{T,phot}}$ is a piecewise linear function that modifies photosynthesis to suboptimal temperatures throughout the day and $F_{\max}$ computes the effect of CO_2 as $F_{\max} = \tau_{CO_2} V_{CO_2}$.

Respiration is described by:

$$V_{\text{res}} = \int_{t_i}^{t_e} Q_{10}^{(X_{\text{T,a}}-20)/10} r_m (X_W - X_{\text{FM}}) \mathrm{d}t \tag{2.88}$$

where Q_{10} is the sensitivity of respiration to temperature and r_m is a maintenance respiration coefficient, t_i is the initial time and t_f is the final time. Breathing is calculated updating in time the matter that has become ripe fruit, this means that once a fruit has reached maturity is immediately harvested.

The dry matter of fruits (X_F), is described by

$$\frac{\mathrm{d}X_F}{\mathrm{d}t} = GR_n\alpha_F f_F(\bar{X}_{\mathrm{Td,a}})\,(1 - \exp(-\nu(X_N - N_{\mathrm{FF}})))\,g(X_{\mathrm{T,day}}) \quad \text{if} X_N > N_{\mathrm{FF}} \tag{2.89}$$

where ν is the vegetative-fruit transition coefficient, α_F is the maximum partitioning of new growth to fruit, N_{FF} is the number of nodes per plant when first fruit appears, function $f_F(\bar{X}_{\mathrm{Td,a}})$ computes the effect of average daily temperature on the distribution between vegetative and reproductive growth under low temperatures, while function $g(X_{\mathrm{T,day}})$ modifies the distribution of fruits in very hot days [211, 212, 339], where $X_{\mathrm{T,day}}$ is the average temperature of the daylight hours during the day.

The dry mater of mature fruits (X_{MF}) is given by:

$$\frac{\mathrm{d}X_{\mathrm{MF}}}{\mathrm{d}t} = D_F(\bar{X}_{\mathrm{Td,a}})(X_{\mathrm{F}} - X_{\mathrm{MF}}) \quad \text{if} \quad X_N > (N_{f1} + K_F) \tag{2.90}$$

where K_F indicates the number of nodes since the first fruit appears until it matures, N_{f1} indicates the number of nodes when the first fruit appears and $D_F(\bar{X}_{\mathrm{Td,a}})$ is a piecewise linear function of average daily temperature [211, 212, 339].

Figure 2.22 shows the behavior of the state variables for the *Tomgro* model when data from the spring season was used. Regarding the dynamic of the state variables, it can be observed an acceptable behavior of the model. In [339], the average absolute error is included, which has the following ranges of values: 0.5 and 1.9 for number of nodes, 0.11 and 0.24 $\mathrm{m}_{\mathrm{leaf}}^{-2}\,\mathrm{m}_{\mathrm{soil}}^{-2}$ for leaf area index (X_{LAI}), 23.9 and 44.4 g m^{-2} for total dry matter (X_W), 19.9 and 42.1 gm^{-2} for dry matter of the fruits (X_F), and 37.1 and 44.9 for dry matter of mature fruits (X_{MF}) when the *Tomgro* model was applied using data from the spring–summer and the fall-winter season, one of which was performed with data from a commercial operation.

2.2.1.4 The Tomsim Model

The *Tomsim* model [182–184] is oriented towards the knowledge of the tomato dynamic beginning at the flowering stage; it is considered no restriction of water and nutrients and an optimum control of pests and diseases. This is a model based on photosynthesis and it allows the user to know with great detail the clusters appearance, the growth of the vegetative segment between two consecutive clusters (called vegetative unit), and the growth of every cluster. Distribution of dry matter is regulated by the sink organs and it is independent from dry matter production [182, 183]. It is required to have data for the 24 h: Temperature, light intensity, and carbon dioxide concentration. Partition of photoassimilates among sink organs occurs every day according to the relative strength of the sink, calculated considering the sum of all the sinks [182–184]. This model utilizes previous work [146] for the simulation of photosynthesis and production of dry matter. The photosynthesis equation used

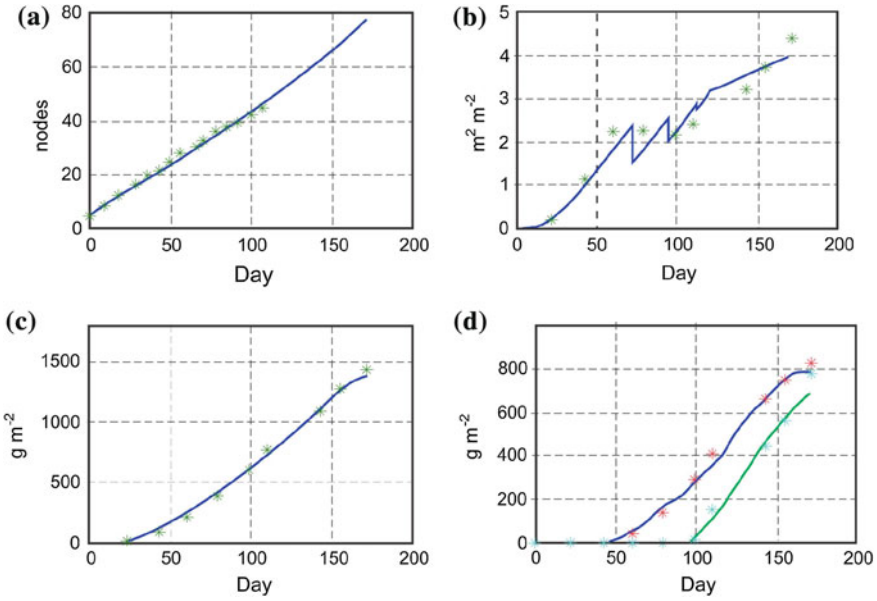

Fig. 2.22 *Tomgro*: Simulation (*solid line*) and observed data (*asterisks*) during a spring season. **a** Number of nodes. **b** LAI. **c** Total dry weight. **d** Dry weight of fruits and mature fruits. Day = day after planting

was validated in different experiments with different ranges of CO_2 concentration and PAR radiation [184].

Photosynthesis in this model is based on the *Sukam* model [146], which considers photosynthesis for leaves and extrapolates the results to the whole canopy. The model includes the calculation of absorption of photosynthetic radiation for the layers of the canopy, the diffused and the direct photosynthetically active radiation that reaches every layer of leaves, and it is calculated in function of temperature, carbon dioxide, and photosynthetically active radiation. The model takes into consideration dark respiration, which is function of temperature, and also maintenance respiration which considers the state of the different organs (leaves, stem, roots, and fruits).

Flowering estimation is described by state X_{NT} describing number of trusses. The function of clusters appearance per day in *Tomsim* has been developed with data from different experiments [184, 227] and it is an empirical function that considers temperature as an essential variable. Number of fruits is not modeled and it is an entry from the user.

Trusses are harvested after a growth period from anthesis to ripening of fruits. This growth period decreases with temperature, but the degree of this reduction is different according to the fruit development [184]. Every day the model estimates the stage of development of the fruit, which has values between 0 and 1. When the development stage of the fruit reaches the value of one, it is ripe and must be harvested.

The potential growth of every cluster is the maximum growth achieved under optimum conditions. This is modeled from different parameters obtained in an empirical way. The vegetative unit is formed by the stem section and three leaves between two clusters, although the number of leaves before the first cluster is between 9 and 12, and therefore the first vegetative unit is assumed to be 2.5 times bigger than the other [182, 184].

This model considers the detailed growth of the vegetative part of the crop, and it measures the growth using vegetative units. The potential growth of the vegetative units is a function of the daily average temperature and the potential growth of one fruit in the cluster of reference. A vegetative unit starts its growth approximately 3 weeks before the corresponding cluster.

The allocation of dry matter is regulated by the sinks or destination of assimilates, and these sinks are the clusters and the vegetative units. The available assimilates (g of CH_2O per plant) are distributed among the total number of sinks per plant according to the strength of every sink, which is the potential growth of clusters and vegetative units. The sink strength for roots is established in 15 % of the vegetative sink strength. The distribution within the vegetative part of the plant is 7:3:1.5 for leaves, stem, and roots, respectively [183]. If the amount of available photoassimilates is equal or greater than the sum of the sinks, every organ will grow to its potential and the unused photoassimilates will be sent to storage. The next day these reserves are added to the newly formed photoassimilates [182, 184].

LAI is simulated based on the estimated dry matter of the leaf area X_{LDW} and the specific leaf area (SLA). The model considers that leaves are eliminated when their corresponding cluster reaches a development stage of 0.9 [183]. Specific leaf area (SLA) expressed in $\mathrm{cm}^{-2}\,\mathrm{g}^{-1}$ is modeled by *Tomsim* with a function of day of the year.

Figure 2.23 shows the simulations and the observed data for a spring-summer season applying *Tomsim*. With the exception of the regression coefficient for the exponential equation that relates relative growth rate and maintenance respiration which value is 10, and the extinction coefficient for diffuse radiation which value is 0.712 [339], the values of the utilized parameters are in [182] and [184]. This model was applied with data from different spring-summer and fall-winter seasons, and the average errors in relation to the observed data were the following: 0.41–0.59 on number of trusses (X_{NT}), 21.6–52.3 gm^{-2} on total dry matter (X_W), 9.9–28.3 gm^{-2} on dry weight of fruits (X_F), 1.53–3.21 gm^{-2} on the first cluster, 2.07–4.75 gm^{-2} on the second cluster, 1.57–2.73 gm^{-2} on the third cluster, 2.61–6.24 gm^{-2} on the fourth cluster, and 1.68–6.46 gm^{-2} on the fifth cluster [339, 341]. Equations of the model and values of parameters are not included for the sake of space, but can be found in [182, 184, 339, 341].

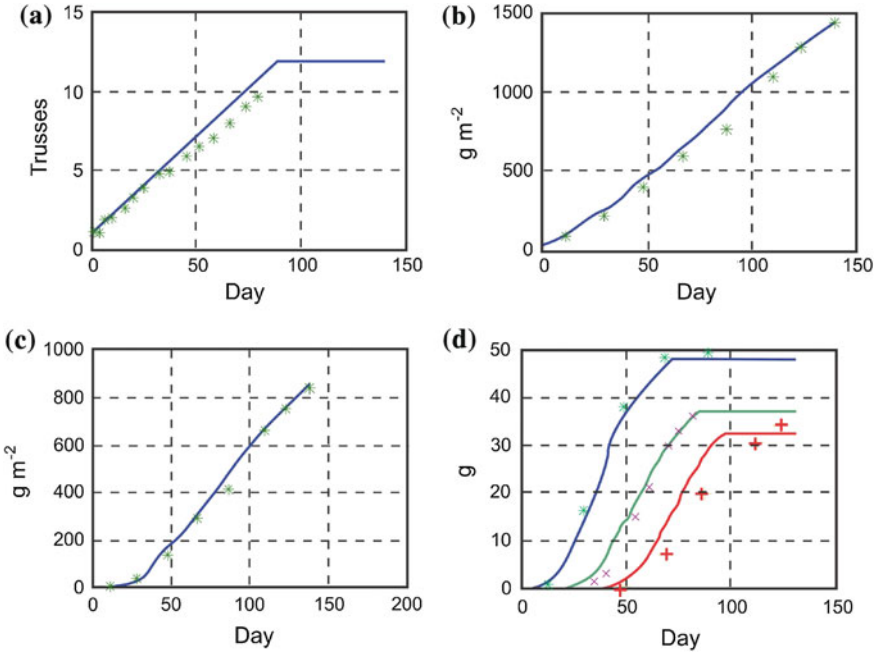

Fig. 2.23 *Tomsim*: Simulation (*solid line*) and observed data (*asterisks*) during a spring season crop. **a** Amount of clusters. **b** Total dry weight. **c** Dry weight of fruits. **d** Dry weight of clusters 1, 3 and 5

2.2.2 Effect of Salinity, Water Deficit and Vapor Pressure Deficit in Yield

2.2.2.1 Salinity

The growth models of crops are oriented to the evaluation of the production of dry matter; however, the product that growers take to the market is a fresh product, with water content between 93 and 95 % [186]. Simulation models allow the users to estimate fresh yield by knowing the relationship dry weight/fresh weight of fruits. According to some reports, the content of dry matter for tomato mature fruits is 5 % in fall and 5.6 % in spring for the southeast of Spain [33, 339].

It is known that salinity in the root media causes a yield reduction of the commercial fresh weight of fruit vegetables [92, 120, 250, 411, 413, 414, 418]. Table 2.13 indicates the magnitude of the decrease of yield for different salinity degrees, expressed as electrical conductivity (EC [mS cm^{-1}], represented by state variable X_{EC}) of the nutrient solution. In some conditions there was no effect, for example in [411] it is indicated that a beefsteak variety grown in the fall had a solution that lasted 90 days and had EC of 2.9. EC was then increased to 5.0 for one treatment and to 6.8 for another, and these values were applied from day 91 to day 130; as a result, there

Table 2.13 EC indicated by different authors and related to yield of tomato fruits

Crop	Length of the season (months)	Threshold value (mS cm^{-1})	Decrease of commercial yield (% per unit of X_{EC})	Min and max X_{EC} in experimental trial (mS cm^{-1})	Reference
Estafette	4 (autumn)	2.5	2.3	2.5–5.2	[413]
Turbo	8 (spring–summer)	2.9	7.2	2.5–5.2	
Abunda,	4 (autumn)	2.4	5.2	2.4–4.6	[411]
Calypso,	4 (spring)	2.6	7.0	2.6–3.5	
Angela					
Rambo,	5 (winter–spring)	2.7	9.8	2.7–13.0	[415]
Daniela,					
Moneymaker					
Daniela	5 (spring)	3.79	8.7	2.72–7.84	[261]
Chaser	5 (spring VPD=0.49)	2.0	5.1	2.2–9.3	[250]
	5 (spring VPD=0.30)	2.0	3.4	2.2–9.0	
nd	8 (autumn–winter –spring)	3.4	4.4	3.4–5.7	[418]
Gokce F1	4 (spring)	1.9	8.3	2.8–6.2	[120]
FA 361	6 (autumn)	1.9	9.1	2.3–5.8	
Counter	3 (spring)	3.0	5.7	1.0–11.0*	[92]
	3 (autumn)	3.0	1.5	1.0–11.0*	
Capello	6 (autumn)	2.5	3.7	2.5–5.5	[479]
L1	4 (spring)	1.8	9.5	1.9–9.1	[346]

Done with information from the cited sources

were no significant differences in yield when compared with 2.9 during the whole season treatment. By contrast, when the increase in EC was done at day 60 there was a decrease of 3.1 % for every mS cm^{-1} increased.

It has been also reported that under poor conditions of light and during early stages of growth, high salinity values did not affect yield in the long run [411]. The same was observed when high values of salinity and poor light conditions were applied in the reproductive stage. In some experiments where EC ranged from 2.0 to 5.6 dS m^{-1} there was no yield reduction in commercial fruits developed under 100 μmol m^{-2} s^{-1} of artificial light, a CO_2 concentration of 800 ± 200 ppm and temperatures between 17 and 21 °C [112].

When the concentration of nutrients in the nutrient solution is lower than crop requirements the yield decreases [411]. Figure 2.24 describes the relationship between EC in the rhizosphere and relative yield; if the EC is below a minimum or above a maximum there is a decrease in yield [411]. The maximum value above which yield decreases is estimated at 2.55 mS cm^{-1} and it is the average of the values in Table 2.13. The average reduction in yield is 6.1 % per mS cm^{-1}. Figure 2.24

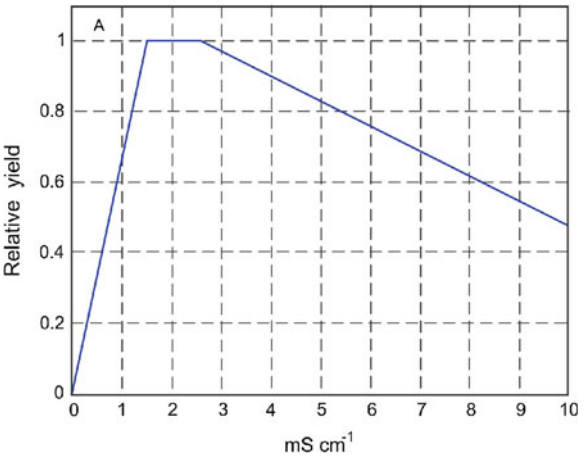

Fig. 2.24 Effect of salinity on tomato yield, relative yield respect to EC in the nutrient solution (mS cm^{-1})

indicates the relationship between relative yield and salinity or nutrients deficiency, which can be indicated as a function in sections.

From the information in Table 2.13, it is possible to formulate a function of yield or fresh weight of fruits which is valid when EC of the nutrient solution is above the threshold of yield:

$$X_{\mathrm{FF}} = \frac{X_{\mathrm{MF}}}{D_{\mathrm{mc}}}[1 - R_y(X_{\mathrm{EC}} - S_t)] \tag{2.91}$$

where X_{FF} is fresh matter of fruits, X_{MF} is dry matter of mature fruits, D_{mc} is the content of dry matter in mature fruits, R_y is the reduction in yield per unit of EC of the nutrient solution X_{EC}, and S_t is the threshold of electric conductivity above which there is a yield decrease.

It should be noted that there is an effect in the decrease of yield when the ionic concentration in the nutrient solution expressed as X_{EC} is below the threshold. Also, it should be noted that yield decrease is greater when data from the experimental trials in Spain are considered, compared to average results dealing with salinity.

2.2.2.2 Water Deficit

The effect of poor supply of water on yield of crops has been extensively studied with the main goal of developing irrigation recommendations [109]. The equation of Stewart, which relates yield with water supply, was evaluated in different crops and is as follows:

$$1 - \frac{V_R}{V_{\text{Rmax}}} = c_{ky}\left(1 - \frac{V_{ET}}{V_{ET_{\text{max}}}}\right) \tag{2.92}$$

where V_R is yield obtained with limited irrigation, estimated in function of actual evapotranspiration (V_{ET}), V_{Rmax} is the yield obtained in non-limited irrigation conditions, equivalent to the maximum evapotranspiration ($V_{ET_{\text{max}}}$), and c_{ky} is a sensitivity to evapotranspiration deficit factor or a crop response factor [109]. The value of c_{ky} is determined at 0.68 [218] for greenhouse tomato grown in soil with insufficient irrigation. Reference [109] provides values of 1.0 and 1.1 for c_{ky}.

Applying a deficit irrigation, a linear equation for fresh fruit yield is developed for greenhouse tomato, in which yield is function of irrigation applied as function of evapotranspiration [37].

$$V_R = 0.99 V_{\text{Rmax}}\left(\frac{V_{ET}}{V_{ET_{\text{max}}}}\right) - 0.14 \tag{2.93}$$

Considering the references mentioned above, and assuming in the growth models the crop is well irrigated with maximum evapotranspiration, it is possible to develop simulation models to show the effect of less amount of water than the required.

2.2.2.3 Vapor Pressure Deficit

Vapor pressure deficit (VPD) also has effect on yield, which is important during the winter months in poor ventilated greenhouses or during the hottest months of the year. When studying the effect of high VPD, [44] compared 2.2 kPa and 1.6 kPa calculated during the six driest hours of the day and found a decrease of 16 % in tomato yield with high pressure deficit. At the other extreme, a yield decrease of approximately 30 % was reported when the treatments of 0.5 kPa (control) and 0.1 kPa (high humidity) were compared and estimation was done considering the average of the 24 h of the day [286]. A piecewise linear function was created with the mentioned data, as shown in Fig. 2.25.

2.3 Water Models in Artificial Substrates

2.3.1 Water Dynamics

Water is important for plants; it is a constituent of vegetable tissues, a solvent, a reagent, keeps cellular turgor [233, 276], and is an excellent medium for temperature regulation [276]. Between 80 and 90 % of fresh weight of plants is water. A decrease in water content is accompanied with a loss of turgor and wilting, cellular elongation halt, stomatal closure, reduction of the photosynthetic activity, and malfunctioning of

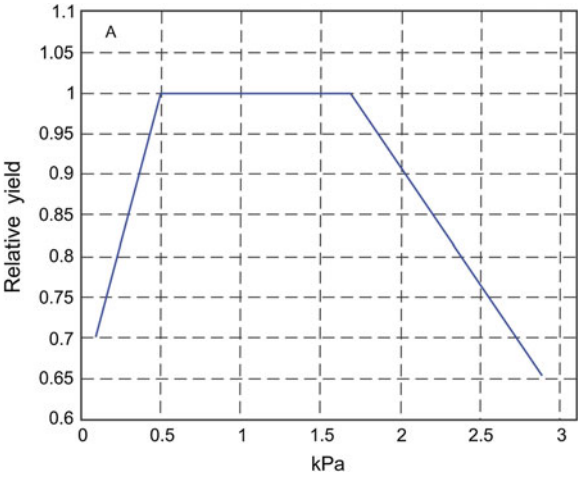

Fig. 2.25 Relative yield respect VPD [kPa]

many other basic metabolic processes [233, 434]. A nonadequate supply of water to maintain turgor results in immediate reduction of vegetative growth [233, 276, 434].

It is important to analyze the hydric balance in a crop because this analysis helps to understand the medium-plant-atmosphere continuum dynamics and makes it possible to efficiently manage water. This analysis can be done considering: Water in the substrate, water uptake, transport from root to leaves, and movement from leaves to the atmosphere. Since there is a close relationship between the processes mentioned above, this division is only for analysis purposes.

2.3.1.1 Dynamics of Water in the Substrate

Movement of water is studied and explained in function of the energetic state of water, describing the flux through a substrate [68] or the soil [233]. In this type of flux it is considered that inertia in the movement of water is small, and what is important is the potential energy [68].

It is possible to define a function for hydric potential that takes values in all the points of the substrate, in which the flux of water at every point goes from greater to smaller potentials and the direction is determined by the maximum variation of the potential [68], although [233] indicates that there are situations where this tendency is not followed. The speed of water is proportional to the gradient of such potential [68], and the proportionality constant is a specific characteristic for the substrate utilized. In other words, the same gradient of potential will generate a different flux in different substrates [68].

The potential energy per unit of mass (or volume) of water is the hydric potential, and at any point of the substrate it receives various contributions due to several factors as gravity field, influence of dissolved ions, and local pressure [68, 233].

Therefore, the hydric potential of the substrate ψ_{hs} is the sum of potentials that can be measured in an independent way [68] and can be expressed as

$$\psi_{hs} = \psi_{pr} + \psi_g + \psi_{os} \tag{2.94}$$

where ψ_{pr} is the potential of pressure, ψ_g is the gravitational potential and ψ_{os} is the osmotic potential. The potential of pressure is composed of the matric potential ψ_m, the pneumatic potential ψ_n and the enclosing potential ψ_e [68]. The gravitational potential can be expressed in terms of height differences between the considered and the reference point:

$$\psi_g = c_{\text{den,w}} c_g c_z \tag{2.95}$$

where $c_{\text{den,w}}$ is density of water, c_g is acceleration of gravity and c_z is the level in relation to the reference point. The osmotic potential is given by osmotic pressure of the substrate solution, and it can be determined with the equation of the state of the perfect gases:

$$\psi_{os} = -c_{\text{csol}} c_R T_s \tag{2.96}$$

where c_{csol} is the solutes concentration, c_R is the universal constant for gases, and T_s is the absolute temperature. The potential of pressure depends on the local content of water and for practical purposes the matric potential is considered as the only component of the potential of pressure [68]. The matric potential can be understood as the suction force applied by the plant to extract the water retained by the substrate [10, 223].

On the other hand, and from the point of view of substrate characterization, one important aspect is the capacity for water retention in function of its physic characteristics, porosity, structure, size, and distribution of size of particles [10, 68]. Substrates with particles between 1 and 10 mm have little variation in the amount of retained water, and the capacity of water retention increases when particles are smaller than 1 mm [10]. The maximum content of water of a substrate is known as container capacity and it is a function of the substrate characteristics and height of the container [10, 68].

Water retained by the substrate expressed as its humidity content follows a nonlinear relationship with the matric potential and shows a hysteresis phenomena; in other words, the humidity content is different if the substrate is getting dry than if is getting wet [68]. The knowledge of this relationship has been the object of many studies [68, 409] and is useful to formulate models of hydric balance.

2.3.2 Water Uptake by the Plant

2.3.2.1 Water Flow Towards the Root

The most important structures in the root system are epidermis and root hairs because they make direct contact with the soil and are the surfaces through which most water and minerals enter the root. A root generally has access to all the available water in a ratio of approximately 6 mm. When the soil (substrate) dries off due to the effect of the matric forces, the movement of water slows down [241].

Water uptake occurs because of potential gradients from the substrate to the roots; there are two uptake mechanisms, active and passive. The first mechanism, also called osmotic uptake, occurs in plants with slow transpiration where the roots behave as osmometers [233], and not as resistances, so in absence of transpiration the uptake of ions to the xylem produces a flow of osmotic nature and therefore it produces a pressure at the root level [426]. Passive uptake in plants happens when transpiration is high and water is suctioned toward the roots [233]. There is evidence that forces involved in the uptake of water to roots (passive process) are caused by a tension created by transpiration of the canopy, which expands to the root xylem [241, 426], although some authors recognize that the relationship between time of response and transpiration strategy of plants is not clear [121].

By definition, and for modeling purposes, the hydraulic properties of roots have two parameters: The minimum gradient of hydric potential to induce a flow, and the hydric conductivity [347]. The flow of the soil solution from soil to root and from root to canopy occurs through a complex structure with variable hydraulic resistances, some of which can be considered serially (in different tissues of the root cylinder) or in parallel (different cellular ways for water) [427].

2.3.2.2 Water Potential in Root and Leaves

In a similar way as the hydric potential in the substrate, in vegetable cells there is a hydric potential determined by the potential of pressure, the osmotic pressure due to solutes and the matric potential, although the matric potential is very small [451]. Measurement of hydric potential in roots is difficult and the commonly used methods are not completely appropriate because they utilize cut roots and the conditions are different from the conditions in roots of intact plants [491]. Different studies of hydric potential in roots indicate that this is influenced by climatic factors and by the plant itself. Hydric potential in roots, specifically the osmotic potential of sap in roots, decreases with flood treatments and the cause is osmolality [204]. This additional osmotic force explains a greater speed of the flow of the sap through the tomato roots of flooded plants, compared with well-drained plants in similar experimental conditions of pressure applied [204].

When severe hydric deficit was applied it was observed a decrease in the hydric potential in leaves, reaching values of -2.5 MPa, although this hydric poten-

tial was partially reestablished when the deficit stopped (−1.8 MPa) in Populus seedlings [398]; in tomato the hydric potential was −1.65 and −0.40 MPa for the maximum hydric stress and when irrigation was restored, respectively [456].

Salinity induces a decrease of hydric potential within the plant (usually measured in the leaves) [78, 266, 364, 375], and the response is different according to the species and the salinity degree.

The minimum gradient of hydric potential for a flow to happen within the roots is in the range of 0.08 and 0.49 MPa. These values were determined in intact plants; the gradient is associated to the presence of exodermis in the root, and it is not correlated to the cortex thickness or with the root diameter [374].

2.3.2.3 Hydraulic Conductivity in the Roots

Hydraulic conductivity is a property of roots and expresses the relationship between water flow and the hydric potential gradient [347]. Hydraulic conductance or resistance in roots has a different magnitude if it is a radial flow through the root cylinder or an axial flow along the xylem, and the axial conductance is smaller [427]. The behavior of conductance or resistance depends on the age of the root and can be different according to the external conditions (salinity or water deficit) or internal factors (nutrition or water needs of the plant) [426].

According to some authors hydraulic resistance in the plant is independent of the flow of water for transpiration, and there is a linear relationship between transpiration and gradients of hydric potential of the nutrient solution in relation to stem, and also of the stem in relation to the leaves [248].

Hydraulic conductivity increases as transpiration speed increases [210, 427], whereas such conductivity decreases when plants are flooded at the beginning of the morning when daylight starts, and this could be originated by the decrease of O_2 in the root zone [108].

2.3.3 Transpiration

Transpiration has been described in Sect. 2.1.1.6, Eqs. (2.42)–(2.47), following the approaches based in the Penman-Monteith equation [241, 301, 419, 451]. Transpiration causes a decrease in the leaf cells hydric potential, and it originates water demand toward the evaporation surfaces. This flow continues as long as there are gradients, which are established step-by-step through the soil–plant system, and they define the flow speed of the water to the leaf [233, 491].

The elements to simulate the movement of water dynamics from the soil or the substrate to the atmosphere, passing by the plant, are: Hydric potentials, resistance of the substrate to water movement, resistance of the roots, resistance of the transfer elements, resistance for the water to leave toward the atmosphere and the architecture of the root system.

2.3.4 Integrated Water Model

In this section a generic model that considers water balance from an integral point of view is presented as a submodel to make connection with an explanatory growth model for the crop and the ecosystem. It is dynamic, explanatory and simple. It considers the amount of water in the substrate, the root and the shoot (leaves, stems and fruits). The model of water balance is combined to a growth model in which the dry matter is divided into structural biomass and nonstructural biomass (storage), whereas the soil or substrate has only one layer. The state variables are mass of water in the substrate, mass of water in root, and mass of water in shoot; it has 30 parameters, 6 of which can be changed [452] to adjust the model to different conditions.

The main dynamics of the state variables, that are mass of water in the shoot (X_{wc}) and root (X_{wr}), are defined by the next equations.

$$\frac{\mathrm{d}X_{wc}}{\mathrm{d}t} = F_{w_{r-c}} - V_{ET} \tag{2.97}$$

$$\frac{\mathrm{d}X_{wr}}{\mathrm{d}t} = F_{w_{s-r}} - F_{w_{r-c}} \tag{2.98}$$

where $F_{w_{r-c}}$ is the water flow from root to shoot, V_{ET} is the flow from shoot to atmosphere, and $F_{w_{s-r}}$ from substrate or soil to root.

When the relative water content in the substrate (θ_r) is greater or equal to field capacity or container capacity (θ_{mx}) the water flow is an excess flow or drainage ($F_{w_{dr}}$) and therefore the mass of water in the soil (X_{wss}) does not change; in other cases it happens that:

$$\frac{\mathrm{d}X_{wss}}{\mathrm{d}t} = F_{w_s} \tag{2.99}$$

where F_{w_s} is the water flow in the soil or substrate. Flows of water are determined according to:

$$F_{w_s} = F_r - F_{w_{s-r}} \tag{2.100}$$

$$F_{w_{s-r}} = \frac{\psi_{hs} - \psi_{hr}}{r_{wsr}} \tag{2.101}$$

$$F_{w_{r-c}} = g_{wrc}\,(\psi_{hr} - \psi_{hc}) \tag{2.102}$$

$$F_{w_{dr}} = F_{w_s} \tag{2.103}$$

where F_r is the irrigation supplied, ψ_{hs}, ψ_{hr} and ψ_{hc} are water potentials of soil, root and canopy, respectively; r_{wsr} is the resistance to the flow from soil to root, and g_{wrc} is conductivity of the flow from root to shoot.

The model is oriented toward the application in intensive crops which have low volume of substrate and the height of the container is less than 15 cm. Under these conditions the hydric potential of the substrate considers Eq. (2.94), in which the gravity potential is negligible because of low height of container. It is also considered that the matric potential (ψ_m) is the most important component of the pressure potential, therefore the hydric potential of the substrate can be estimated with the following equation:

$$\psi_{hs} = (\psi_{os} + \psi_m) \tag{2.104}$$

Taking into consideration the relationship between the matric potential and the characteristic curve for water retention, the expression proposed by van Genutchen [141] is used:

$$S_e = \frac{1}{(1 + |c_{w1}\psi_m|^{c_{w2}})^{c_{w3}}} \tag{2.105}$$

where S_e is the effective water content of the substrate or effective saturation, c_{w1}, c_{w2} and c_{w3} are parameters of shape from the water retention curve. Hydraulic conductivity of the substrate is calculated according to equation [285]:

$$K_{rSe} = S_e^{0.5}\left(1 - \left(1 - S_e^{\frac{1}{C_{w3}}}\right)^{C_{w3}}\right)^2 \tag{2.106}$$

where K_{rSe} is the relative hydraulic conductivity, $K_{rSe} = K_{Se}/K_s$, where K_s is the hydraulic conductivity at saturation.

The potentials in root (ψ_{hr}) and canopy (ψ_{hc}) are defined by the osmotic potential (ψ_{osr} and ψ_{osc}) and the pressure potential (ψ_{prr}, ψ_{prc}) in each, according to the following equations:

$$\psi_{hr} = \psi_{osr} + \psi_{prr} \tag{2.107}$$

$$\psi_{hc} = \psi_{osc} + \psi_{prc} \tag{2.108}$$

$$\psi_{osr} = -\frac{c_R(X_{T,a} + 273.15)\, f_{neor} M_{ner}}{\mu_S X_{wr}} \tag{2.109}$$

$$\psi_{prr} = \frac{c_\varepsilon\left(\frac{c_{prr} X_{wr}}{M_{er}} - 1\right)}{c_{\text{den},w}} \tag{2.110}$$

where c_R is the universal constant of gases, M_{ner} is the nonstructural root dry matter, M_{er} is the structural root dry matter, X_{wr} is the mass of water in roots, $X_{T,a}$ is the air temperature, f_{neor} is the osmotically active storage fraction of the root, μ_S is the molal mass of storage, c_{prr} is a parameter that affects the pressure component of the hydric potential within the root, c_ε is the parameter of rigidity of the cell wall and

$c_{\text{den},w}$ is water density. In a similar fashion are defined the osmotic potential and the pressure potential for the canopy.

The resistance to the passage of water from the soil to the root (r_{wsr}) and hydraulic conductivity (g_{wrc}) from the root to the canopy, are defined as

$$r_{wsr} = \frac{c_{sor} V_{\text{den},r}}{K_{so} M_{er}} + \frac{c_{rsr}}{V_{\text{den},r}} \left(\frac{M_{er} + c_{kwrs}}{M_{er}} \right) \tag{2.111}$$

$$g_{wrc} = c_{nw} \frac{X_{wr} X_{wc}}{X_{wr} + X_{wc}} \tag{2.112}$$

where c_{sor}, c_{rsr}, c_{nw}, and c_{kwrs} are parameters affecting the resistance between soil and root, K_{so} is the soil hydraulic conductivity and $V_{\text{den},r}$ is the density of roots.

It is known that root hydraulic resistance is variable, and that with high transpiration amounts such resistance lowers leading to a quick uptake of water [365, 427]. A coefficient is included (c_{khr}) [-] in which the resistance to water flux from soil to root is modified (r_{wsrm}) as an effect of transpiration through an exponential function:

$$r_{wsrm} = r_{wsr} (\exp(-c_{khr} V_{ET})) \tag{2.113}$$

With the aim of adapting the model, some parameters can be considered as just one. The parameters are the nonstructural osmotically active fraction of dry matter for root and canopy (f_{neor}, f_{neoc}) and the molal storage mass (μ_S). The relationship between these parameters is indicated by the following expressions: $c_{far} = f_{neor}/\mu_S$ and $c_{fac} = f_{neoc}/\mu_S$.

The model was adapted to the *Tomgro* model mentioned before, which has been adjusted to estimate structural and nonstructural dry matter utilizing the results of [331] for the tomato crop.

Figure 2.26 shows the dynamic of the water potential in the substrate, root and canopy during 4 days with data of the integrated water model. In the middle of the day the plants have the most negative water potential, therefore it is the period when the crop is more susceptible to stress in the case of having insufficient supply of water in the substrate. By contrast, the less negative water potentials are at dawn, in this moment the water in the system substrate-root-canopy can be in equilibrium.

Figure 2.27 shows the behavior for the water content in the substrate, the dynamic of the simulated and measured data can be seen in this graph.

2.4 Disturbance Forecast

Automatic weather forecasts are important to devise control strategies for greenhouses, being necessary to perform long-term (days) and short-term (min) predictions which help the user to obtain optimal trajectories for the controlled variables

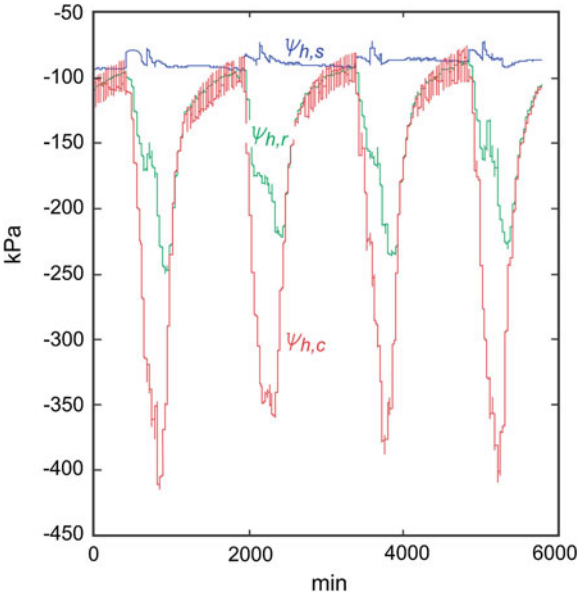

Fig. 2.26 Water potentials in: Substrate (ψ_{hs}), root (ψ_{hr}) and canopy (ψ_{hc})

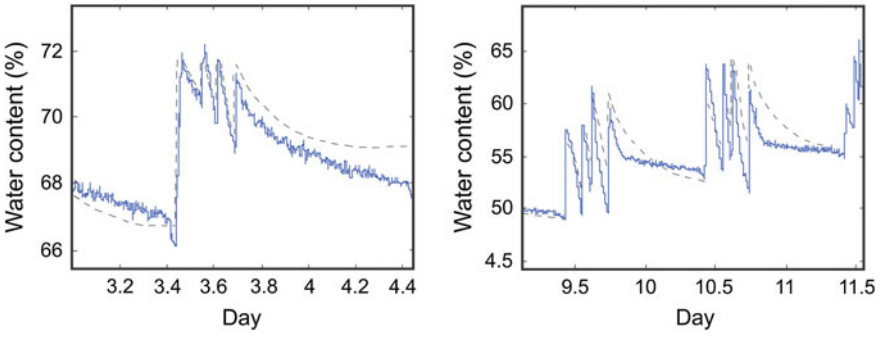

Fig. 2.27 Water content simulated (*continuous*) or measured (*dashed*) in the rock wool substrate

taking the desired objectives into account [71, 343]. There are several methods to perform weather forecasts, and can be classified as a function of:

1. Prediction horizon: Long-term [270] or short-term [345] predictions.
2. Used methodology: There exist in literature different methodologies which can be used to estimate disturbances [307, 345, 484]. In this section, three different methods are summarized: (i) Pattern search models, (ii) Time-series models and, (iii) Artificial Neural Networks (ANN).

2.4.1 Pattern Search Based on the Information Provided by the AEMET

The AEMET which depends on the Agriculture, Food and Environment Spanish Ministry is the reference organism regarding weather forecasts in Spain. These forecasts are obtained through the execution of a limited area numeric prediction model (HIRLAM) based on the environmental conditions provided by the model of the European Centre of Weather Forecasting Medium Range (ECWFM). Therefore, the weather forecasts daily updated (from 05:45—UTC) are published in the Web of the AEMET (http://www.aemet.es/es/portada) and they offer a prediction horizon equal to 7 days for each city of Spain, and equal to 4 days for each province and autonomous community. In Fig. 2.28 a snapshot of the information provided by AEMET by city is shown. More specifically, this information mainly comprises sky state (sunny, cloudy, rainy, etc.), the probability of precipitation, maximum and

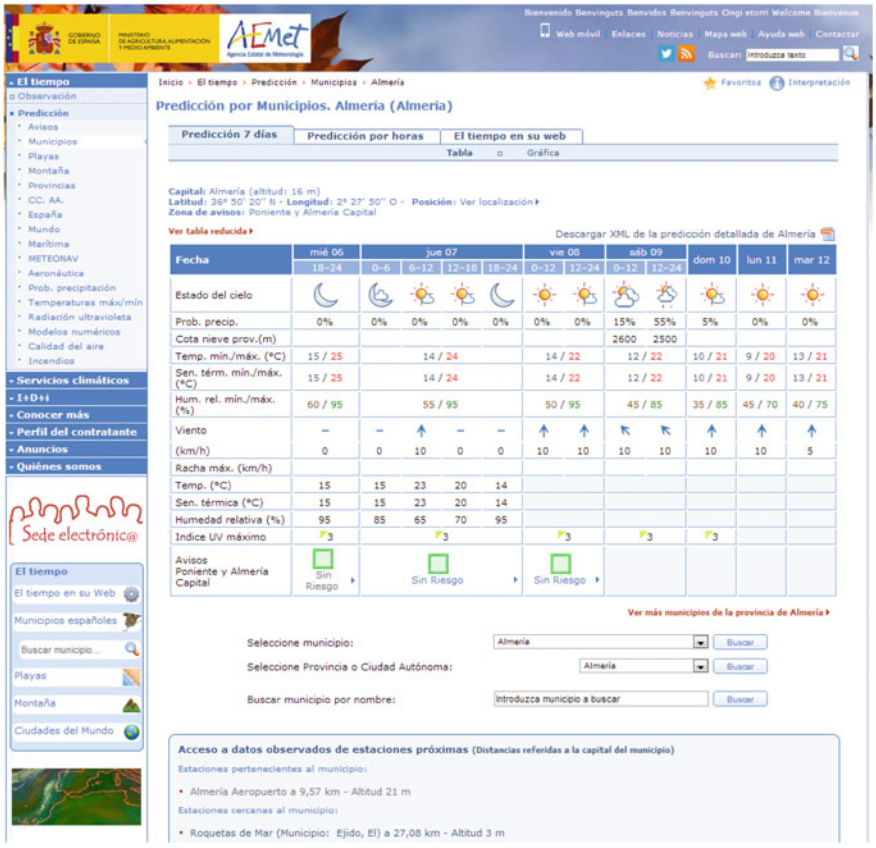

Fig. 2.28 Weather forecast by city. *Source* http://www.aemet.es/es/portada

minimum expected temperature, thermal sensation and relative humidity, wind velocity and direction, and the maximum ultraviolet index.

Moreover, as environmental variables usually repeat each year certain behavior patterns, it is possible to use historic data series as weather forecast. In [303] a predictor module, whose main architecture can be observed in Fig. 2.29a, which provides accurate short/long-term weather forecasts of these outdoor variables which affect the indoor climate of a closed environment is shown. More specifically, the predictor module is able to integrate the information provided by AEMET with data series obtained since 1994, and look for the best pattern equivalent to the forecasts performed by AEMET. For this, the predictor module follows the general algorithm which is shown in Fig. 2.29b. Therefore, it is able to interpret the predictions provided by AEMET, and perform a pattern search within the historic database with the main objective of obtaining a set of patterns with characteristics similar to the predictions of AEMET. Then, each one of the patterns are analyzed as a function of several constraints, as the length of the prediction horizon and the maximum error allowed, and finally, the best pattern is selected from that ones which fulfill the established constraints. If any of the patterns satisfies these constraints, they are relaxed, and the pattern search process starts again.

In order to test the performance of the proposed methodology, to obtain short- and long-term predictions of two important environmental variables, outdoor irradiance and temperature, have been used. On the one hand, for short-term predictions a prediction horizon equal to one day is fixed and, on the other hand, for long-term predictions the prediction horizon is between 60 and 90 days. As can be observed in Figs. 2.30 and 2.31, this methodology provides acceptable results for both long-term and short-term predictions.

2.4.2 Time-Series Models

In general, weather disturbances are represented as time-series structures since they present stochastic behavior. Therefore, a prediction of weather disturbances can be obtained through time-series models. Such models are based on the assumption that the modeled data is autocorrelated and characterized by trends and seasonal variations [323]. Hence, the main objectives derived from time-series methods are: modeling, prediction, and characterization. More specifically, prediction by means of time-series methods require, first, to identify the pattern observed in the data, and second, to propagate it in time with obtained trends and integrate with other data. To do that, in [323] four different well-known time-series methods are analyzed and compared: Discrete Kalman Filter (DKF) [377], discrete Kalman Filter with Data Fusion (DKFDF) [300], Exponentially Weighted Moving Average (EWMA) [464], and Double Exponential Smoothing (DES) [297]. More specifically, they are used to estimate short-term forecasts of solar radiation from real data of a meteorological station placed in the Almería type greenhouse. Nevertheless, the proposed methods and methodology can be easily extrapolated to any location with an appropriate

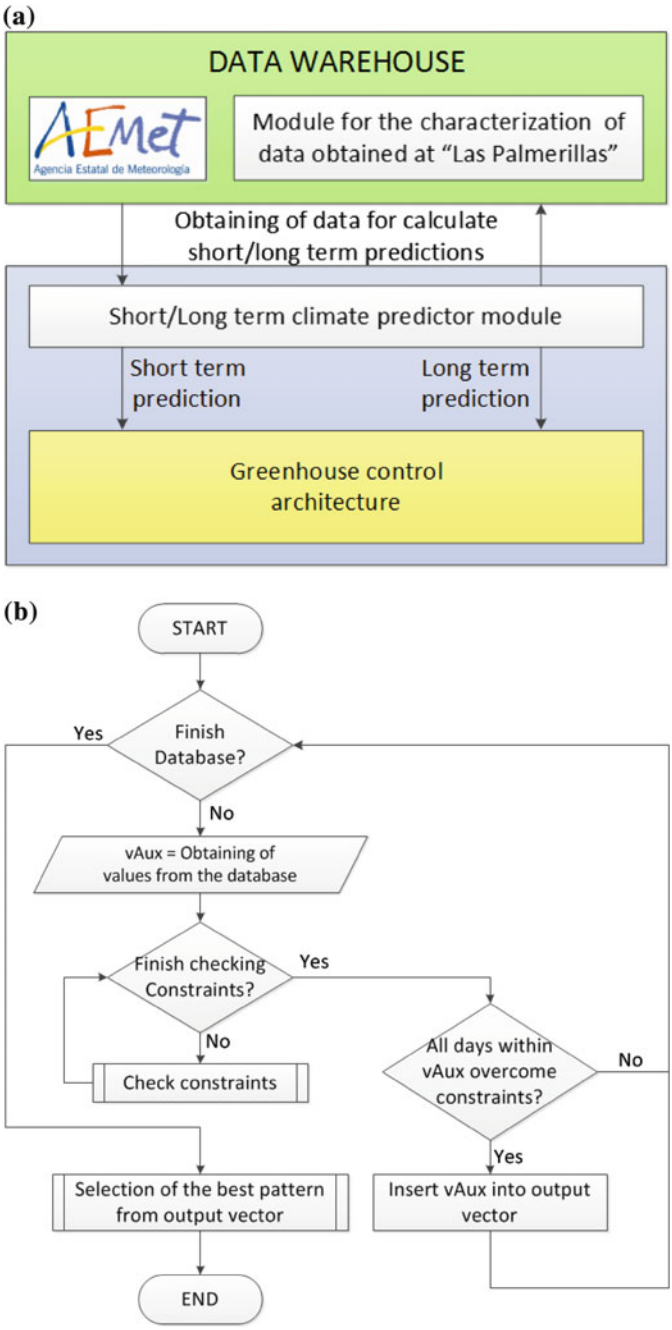

Fig. 2.29 Pattern Search based on the Information provided by the AEMET [303]. **a** Main scheme. **b** General algorithm for the short/long-term climate predictor module

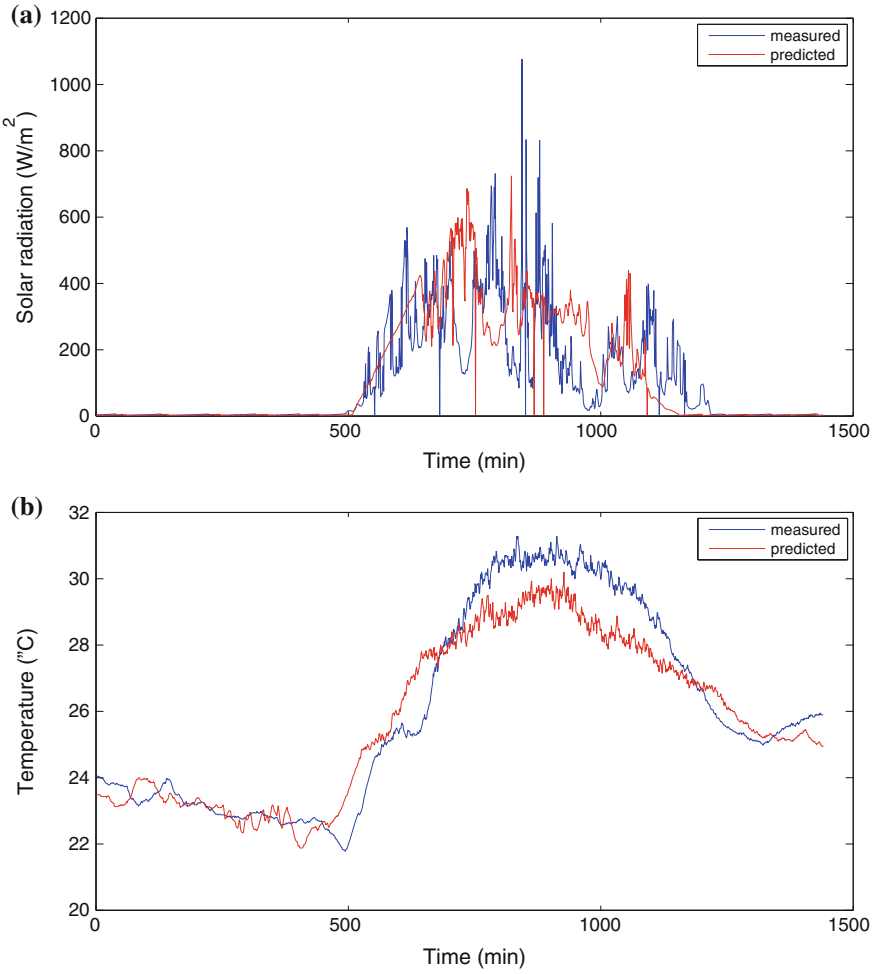

Fig. 2.30 Short-term prediction. *Source* As courtesy of the authors [303]. **a** Short-term prediction of outdoor irradiance. **b** Short-term prediction of outdoor temperature

meteorologic station. A complete description of each of these methods can be found in [323].

To validate the application of the previous forecast methods, one year of meteorological data collected with a sample time, $t_s = 1$ [min] have been used. Figure 2.32 shows the results obtained for each of the time-series methods mentioned previously and for the prediction of solar radiation under different conditions, a clear day and a day with passing clouds. In addition, a short-term horizon equal to 15 samples, that is, 15 min has been used. The obtained results [323] are good since a precise approximation is obtained with all the analyzed methods.

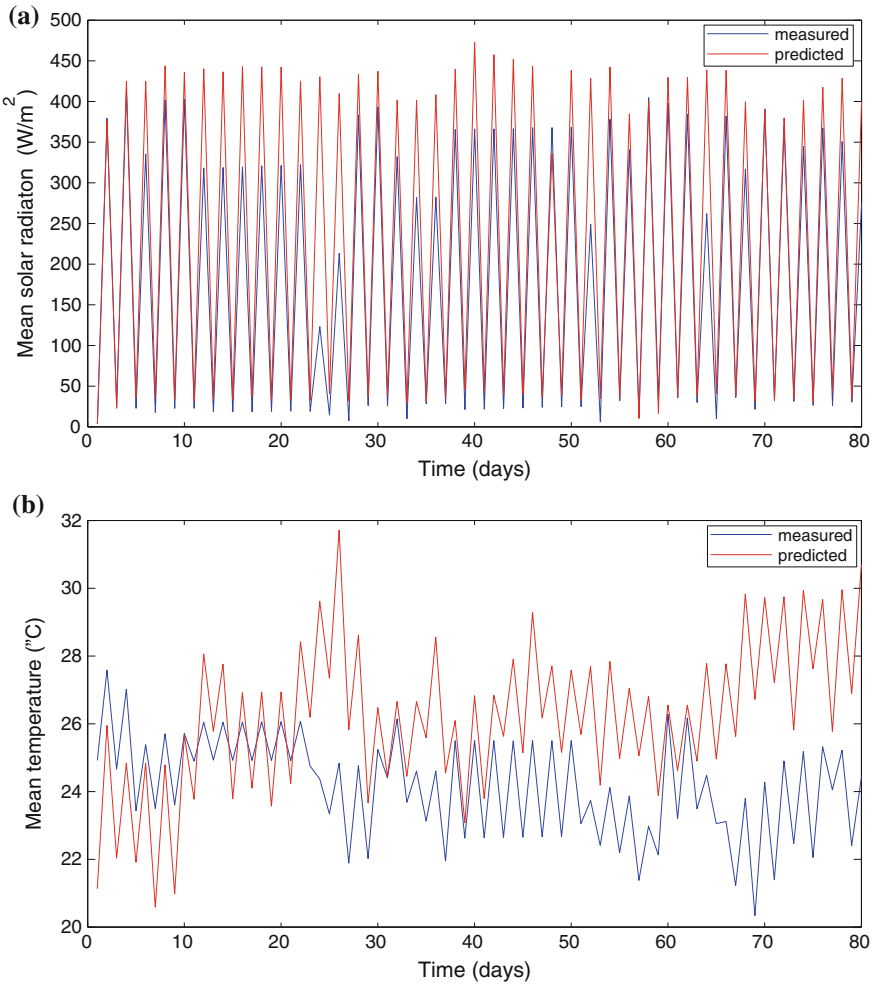

Fig. 2.31 Long-term prediction. *Source* As a courtesy of the authors [303]. **a** Long-term prediction of outdoor irradiance. **b** Long-term prediction of outdoor temperature

2.4.3 Artificial Neural Networks

Finally, ANN can be also used to obtain disturbance models since, as mentioned previously within the climate ANN approximation, their design is based on training and it is not necessary to perform any statistical assumption for the training dataset. As example of the application of this method to estimate disturbance models, two different approximations to obtain solar radiation and outdoor air temperature short-term predictions have been developed. For this, the methodology commented in Sect. 2.1.3.5 of this chapter has been used. More specifically, in this case, two

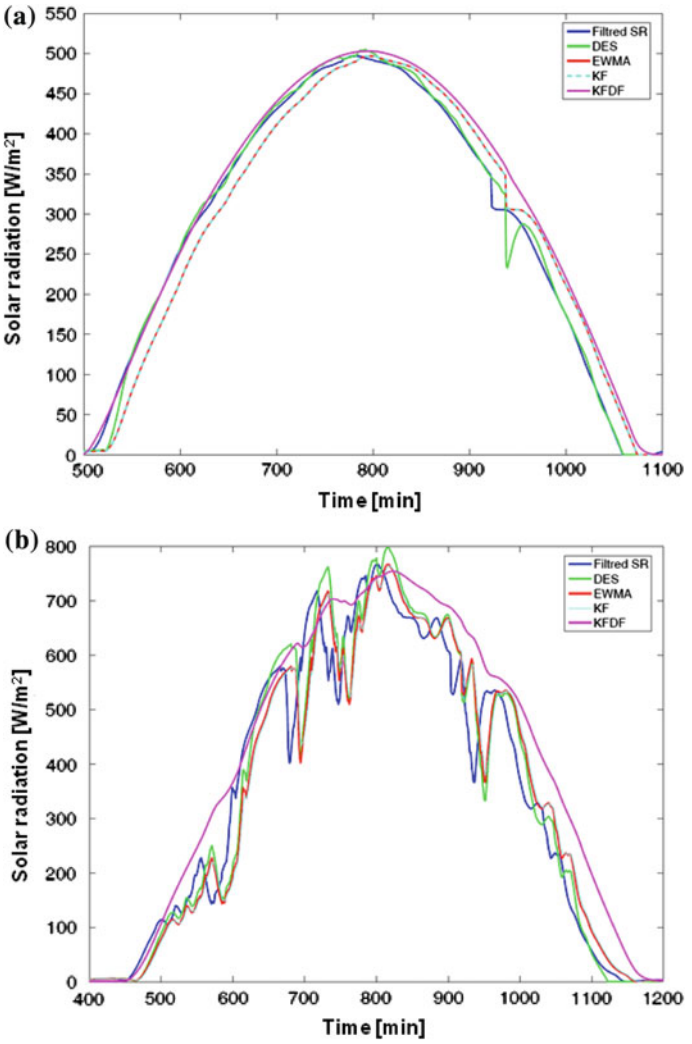

Fig. 2.32 Example of the solar radiation time-series model model under different conditions. *Source* Courtesy of the authors of [323]. **a** Clear day with a 15-sample horizon (15 min). **b** Day with passing clouds and a 15-sample horizon

different NARX ANN have been calculated. The structure of the selected ANN is similar for both, and consists of a NARX configuration with 1 node in the input, a hidden layer with 10 nodes, and a node in the output layer, the solar radiation or the outdoor air temperature prediction. In addition, TDL blocks to take into account past values of the inputs have been included in the ANN architecture. More specifically, a number of past values equal to 4 for each one of the inputs have been used. Besides, different approximation by varying the prediction horizons have been

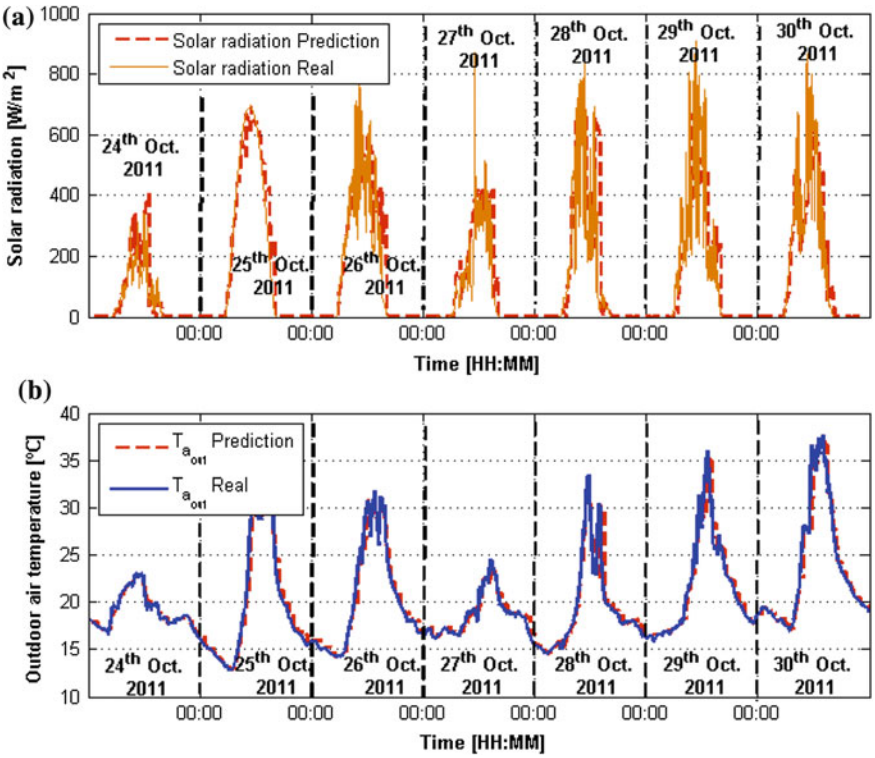

Fig. 2.33 Example of solar radiation and outdoor air temperature ANN short-term prediction models. **a** Solar radiation with a prediction horizon equal to 60 min. **b** Outdoor air temperature with a prediction horizon equal to 60 min

used, $N = [5, 10, 15, 60]\,(\text{min})$. The training process has been performed using a variable-step gradient descent process, namely the Matlab's implementation of the Levenberg-Marquardt algorithm [282]. Furthermore, as training dataset different fragments within the period from 1st September 2010 to 29th February 2012 and a sample time of $t_s = 60\,(\text{s})$ has been used. Finally, the different models were validated using a real dataset from the meteorologic station. More specifically, the selected dataset has a total duration of a week, from 24 to 30th October 2011, and a sample time of $t_s = 60\,(\text{s})$. The results obtained for the solar radiation and outdoor air temperature with a prediction horizon equal to 60 min can be observed in Fig. 2.33a, b respectively.

2.5 Conclusions

A complex nonlinear dynamical model of the greenhouse climate has been developed for the particular conditions of the Southeast of Spain, where the largest concentration of greenhouses in the world is located [355]. These greenhouses are characterized by low-cost structures of medium yield, normally passive or with a low-level of automation and made of plastic cover, taking advantage of favorable outside climatic conditions. In Sect. 2.1.1.2, a description of the dynamic model of the industrial greenhouse climate is formulated. It is composed of six submodels describing the cover temperature, soil surface temperature, first soil layer temperature, inside air temperature and humidity, and PAR radiation onto the canopy. The model implementation is described in detail. It was hierarchically implemented using top-down and bottom-up approaches to provide insight into how the model is organized and how its parts interact. Two different modeling paradigms, block-oriented modeling and object-oriented modeling, were used. The methodology proposed to estimate the unknown parameters of the model is explained based on the fact that the involved physical processes are not coupled. A combination of sequential iterative search and genetic algorithms techniques is used to search the values of the parameters of the model obtaining acceptable results. A sensitivity analysis of the model with respect to the parameters is also included. The model validation process is also explained with different greenhouse structures in winter, spring, and summer seasons, comparing real data measured in greenhouses with data estimated by the model.

The same approach is used for semiphysical models development, aimed at finding simplified models that retain the main nonlinear characteristics of the system but can be used for control purposes. The chapter also includes different structures data-driven models, from linear (based on reaction curve tests or identification), Volterra, and ANN ones.

The second part of the chapter is devoted to develop models for tomato growth for climate conditions of the Southeast of Spain. The tomato crop growth models *Tomsim* and *Tomgro* were calibrated and validated for total dry matter production and calibrated for fruit dry mater production. The parameter estimation was carried out in such a way that the models can be used to simulate the main dynamics of tomato crop growth with differences less than 10 % in total dry matter estimation in both models. The dynamics of tomato crop growth are represented by both models in an acceptable way.

Moreover, water management models are described for soilless systems to supply the adequate quantities without yield reduction but with saving of lixiviates emitted to the environment.

Prediction models for disturbances are also introduced. They play an important role in the hierarchical control architecture where climate setpoints are generated based on the models described in this paper and on predictions of weather and market forecasts.

http://www.springer.com/978-3-319-11133-9

Modeling and Control of Greenhouse Crop Growth

Rodríguez, F.; Berenguel, M.; Guzman, J.L.;
Ramírez-Arias, A.

2015, XXIX, 250 p. 99 illus., 76 illus. in color., Hardcover
ISBN: 978-3-319-11133-9