Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Convex parameter estimator for grey-box models, applied to characterise heat flows in greenhouses

Fjo De Ridder [a,*], Jeroen van Roy [a], Wendy Vanlommel [b], Bart Van Calenberge [c], Maarten Vliex [d], Jonas De Win [c], Bert De Schutter [a], Simon Binnemans [a], Margot De Pauw [a]

[a] KCE (Kenniscentrum Energie), Thomas More University College, Kleinhoefstraat 4, B-2440, Geel, Belgium
[b] Proefcentrum Hoogstraten, Voort 71, 2328, Meerle, Belgium
[c] Proefstation voor de Groenteteelt, Duffelsesteenweg 101, 2860, Sint-Katelijne-Waver, Belgium
[d] Botany B.V., Dr. Droesenweg 7, 5964 NC, Meterik, the Netherlands

ARTICLE INFO

Article history:
Received 4 October 2019
Received in revised form
12 December 2019
Accepted 20 December 2019

Keywords:
Heating
Lighting
Convex optimisation
LED
Greenhouse modelling
Data driven modelling

An algorithm is presented to identify model parameters in grey-box models. The solver is convex, so the global optimum is guaranteed. The method is applied to estimate bulk transfer coefficients in greenhouses from easily monitored data. This model covers the most important processes, such as conduction losses to the environment, heat exchange with neighbouring compartments, heating from the sun and lighting installations, and ventilation losses. Screen positions are also included in the model. Each process is parameterised, so that the specific situation of each greenhouse can be identified. Greenhouse experiments are often repeated in the same greenhouse or performed in parallel. If some model parameters are assumed to remain identical in these experiments, this can be incorporated in the optimiser making it more robust. The estimator is exemplified using measurements from two compartments in a greenhouse; one equipped with LED lighting, the other equipped with HPS lighting. It showed that effective conduction parameters for the greenhouse and screens were similar to values found in literature (5.8 and 5 W m^{-2} K^{-1}, respectively). The model also predicted that both lighting systems provide the same amount of sensible heat at the height of the plants, despite the HPS system consuming 43% more energy. A vertical temperature measurement confirmed that both lighting systems produced the same amount of heat at the height of the plants. The LED system dispersed heat more evenly over height, while the HPS system heated the upper layers more.

* Corresponding author.
E-mail addresses: fjo.deridder@thomasmore.be (F. De Ridder), jeroen.vanroy@thomasmore.be (J. van Roy), Wendy.Vanlommel@proefcentrum.be (W. Vanlommel), bart.van.calenberge@proefstation.be (B. Van Calenberge), maarten.vliex@botany.nl (M. Vliex), jonas.de.win@proefstation.be (J. De Win), bert.deschutter@thomasmore.be (B. De Schutter), simon.binnemans@thomasmore.be (S. Binnemans), margot.depauw@thomasmore.be (M. De Pauw).
https://doi.org/10.1016/j.biosystemseng.2019.12.009

Nomenclature

Indices and sets

$j \in \{1,\ldots,J\}$	Integer, Neighbour index, with J the total number of neighbours
$i \in \{1,\ldots,M\}$	Integer, Heat source/sink index, with M the total number of sources/sinks
$t \in \{1,\ldots,N\}$	Integer, Sample index, with N the total number of samples
$k \varepsilon \Omega$	Integer, Index indicating the common parameters, with Ω the subset of common parameters
$\Omega \subseteq \{1,\ldots,K\}$	Set of integers, Subset of common parameters

Symbols

α_t	$m^3\ m^{-2}\ h^{-1}$, Ventilation rate
ϵ_t^{obs}	−, Observed external variable
Δt	h, Sample period
δ_t^{screen}	Binary, Screen position
δ_t^{vent}	Binary, Window position
θ	−, Model parameters
λ_k^l	−, Lagrange multiplier
ρ	−, Update parameter for Lagrange multiplier
$\chi_t^{exterior}$	$kg\ m^{-3}$, Absolute humidity outside the greenhouse
χ_t^{obs}	$kg\ m^{-3}$, Absolute humidity inside the greenhouse
c_p	$kWh\ kg^{-1}\ K^{-1}$, Specific heat capacity of fluid in heating system
L	$kWh\ kg^{-1}$, Latent evaporation heat of water
$\dot{m}_t$	$m^3\ h^{-1}$, Flow rate of heating system
c	$kWh\ K^{-1}$, Heat capacity
P_t^{light}	kW, Electric power used by the lamps
P_t^{Sun}	$kW\ m^{-2}$, Power delivered by sun
Q^i	$W\ m^{-2}$, Heat for process i
$Q_{t+1/2}^{i,obs}$	$W\ m^{-2}$, Observed heat for process i
S^{ground}	m^2, Common surface with environment
$S^{neighbor\ j}$	m^2, Common surface with neighbour j
$T_t^{exterior}$	K, Observed exterior temperature
$(T_t^{in} - T_t^{out})$	K, Temperature difference between incoming and outgoing water in heating system
T_t^{model}	K, Modelled temperature
$T_t^{neighbor\ j}$	K, Observed temperature of neighbour j
T_t^{obs}	K, Observed interior temperature
V	m^3, Volume of the compartment
$\hat{x}$	−, Estimated value for variable x

1. Introduction and objective

Greenhouse horticulture is an important component of the European agriculture, with substantial land areas covered with greenhouses, with around 178,000 ha in the European Union (Stanghellini, van't Ooster, & Heuvelink, 2019). Plants and crops are grown in greenhouses all year round, which makes greenhouse horticulture energy intensive. In North-Western Europe, natural lighting is insufficient to optimise plant and crop growth during the winter months (Geelen, Voogt, & van Weel, 2018). Therefore, artificial lighting is used. Traditionally, high pressure sodium (HPS) lamps have been used, since these are relatively cheap to produce and maintain. Recently Light-Emitting Diodes (LEDs) have been used to supplement existing HPS lamps to increase production, but in some greenhouses they have replaced HPS lamps entirely. LEDs produce monochromatic light, so a larger fraction of the light produced can be useful for plants. It is generally known that LEDs are more efficient, and thus produce less heat. However this also affects humidity, or more precisely the absolute moisture deficit, and therefore it also influences ventilation in the greenhouse.

In order to quantify the impact of the lighting installations on the climatic conditions in a greenhouse, comparative experiments are required. Such comparisons should be carried out under similar conditions. This can be achieved by using the same greenhouse and altering the applied artificial lighting consecutively. In this case, environmental conditions, such as outside air temperature, wind speed and direction, sun light, etc. will perturb the experiment. It may very well be possible that the influence these uncontrolled factors is far larger than the difference between the lighting installations. An alternative approach is to conduct the experiments in two similar compartments at the same time, so that environmental conditions are identical. However, the internal temperatures may differ, due to different heating conditions, different states of the plants and different humidity, inducing different ventilation rates. Again, these differences may cause greater deviations than the lighting installation itself.

One way to separate the influence of lighting installations from all other heat sources and sinks is to use a model that quantifies all important processes in each compartment. Differences in heating, losses to neighbouring compartments can be corrected for, so that those quantities of interest are isolated and can be compared.

The objective of this paper is to propose a method that can identify the model parameters and states of a grey-box model. This procedure is exemplified on a thermodynamic model for greenhouses. The model consists of simple parametric processes of all main heat flows and takes ventilation into account. The estimator is based on a solver for convex problems and this is exemplified by an experiment where LED and HPS lamps are compared.

It should be noted that the methods discussed in this paper follow the philosophy of data-driven-modelling. The models are a means to holding a dialogue with nature and to extract useful information from noisy measurements. As a consequence, model complexity is dictated by the quality of the observations made.

1.1. Literature review

Here three different types of models that are often used to model greenhouses are described. This is followed by an overview of parameter estimation techniques.

1.1.1. Models

The first group of models reviewed are white-box models. These use only physical parameters from principal physical experiments. These models are used more often when the number of internal states is large, such as in computational fluid dynamics (CFD) models (Boulard, Haxaire, Lamrani, Roy, & Jaffrin, 1999; Nebbali, Roy, & Boulard, 2012; Torreggiani, Bonora, Tassinari, Benni, & Barbaresi, 2016; Yang, Chu, Lan, Tasi, & Wu, 2017). These models provide insight in the physical processes and are often useful e.g. to simulate the behaviour under different design choices. These simulations are faster and cheaper than building and comparing prototypes. However, often, one is interested in estimating the physical properties from a match of a model on observations. This is often very difficult to achieve with white-box models. So, from a practical point of view, white-box models are not always useful, and thus another type of model was investigated here.

At the end of the spectrum are black-box models. These seek a relation between input and output data, without trying to understand this relation (e.g. del Sagrado, Sánchez, Rodríguez, and Berenguel (2016), Frausto, Pieters, and Deltour (2003); He and Ma (2010); Taki, Ajabshirchi, Ranjbar, Rohani, and Matloobi (2016)). Such models are no longer based on physical insights, but purely aim to optimise a prediction. Interpretation of the parameters is no longer applicable, and separation of the physical processes is also no longer possible. Therefore, such model structures are not considered here.

A compromise between both models described above are so-called grey-box models. They are often based on physical laws, like conservation of energy, and are sometimes called energy balance models (e.g. Kimball (1973) or van Henten (1994)). Many applications can be found in literature, such a Mashonjowa, Ronsse, Milford, and Pieters (2013), Pieters and Deltour (1997a, b) or Roy, Boulard, Kittas, and Wang (2002). The advantage of such grey-box models is that the parameters tell us something about the greenhouse under test and this can often be identified from a single experiment with a few sensors. Thus, a grey-box model will be used in this paper.

1.1.2. Estimators

To match a model based on experimental data, an estimator is required. For simple models that are linear or convex in their parameters, parameter estimation is a simple task. However, nonlinearities often appear in these models. When the model is not convex, parameter estimation becomes a difficult task. Parameter estimation methods for such models often have either a low convergence rate (due to the nonlinear nature of the estimation problem) or suffer from high computational burdens (Peifer & Timmer, 2007).

Two approaches to solve this problem can be found in literature. The first set of solutions improve the search engine. Guzmán-Cruz (2013) for example compared a series of estimation techniques to calibrate greenhouse climate models, like genetic algorithms, evolutionary strategies and evolutionary programming. An alternative comparison, using particle swarm algorithms and comparing them with genetic algorithms was given by Yang et al. (2015) and by Akman, Akman, and Schaefer (2018). Under most conditions such search algorithms can provide reasonable parameter values, but one cannot be sure that no better solution exists. In addition, it might take some time before such algorithms converge. Overall, these improvements increase the convergence region and rate, but they cannot guarantee that the optimum found is either the global optimum or just a local optimum.

In a second set of solutions, a reasonable set of initial values for the parameters is searched for, so that the optimisation algorithm can start close to a good solution. Strebel (2013), for example, used the tangent slope and coordinates of a given solution of the ordinary differential equations at randomly selected points in time. Ding and Wu (2013) used a constrained local polynomial regression algorithm to estimate the unknown parameters with the goal to improve the smoothing-based two stage pseudo-least square estimate. Thus, these two papers proposed methods to generate good starting values, but they could not guarantee delivering the global optimum.

Peifer and Timmer (2007) propose to use multiple shooting (Bock & Plitt, 2017) to find a reasonable compromise between the computational burden of finding a global optimum and the process getting stuck in a local minimum. In contrast with single shooting methods, the convergence region of multiple shooting algorithms is much better. In addition, efficient solvers are available, so that computational times can be kept short (Andersson, Gillis, Horn, Rawlings, & Diehl, 2018). Multiple shooting algorithms, however, are a numerical technique. If the ordinary differential equations used are nonlinear in their states, the solution found is still only a local solution.

In contrast to these papers, a different approach is chosen here which ensures that the models used remain convex. This means that the objective function, that measures the distance between model and observations, has one single and unique minimum. So once a solution is found, it is not necessary to doubt if these parameters are sub-optimal. In addition, very efficient algorithms exist, which can solve the search problem in a few seconds. According to our knowledge, none of the above-mentioned papers were able to propose an algorithm that can track the global optimum.

2. Case studies: trial set-ups and locations

Two experiments were used in this paper. In the first experiment two compartments were monitored for one month. In the second experiments, vertical temperature gradients were measured during night in both compartments.

Each experiment was conducted using two identical compartments of a semi-commercial greenhouse, oriented NE-SW in the Research Station for Vegetable Production, Sint-Katelijne-Waver, Belgium. (51° North). The first compartment was equipped with LED lamps and the second with HPS lamps. Throughout this paper the compartments will be referred to as the LED and HPS compartments, respectively. Each compartment of the greenhouse was 20 m long and 8 m wide. The height of the gutter was 6 m; the ridge was about 1 m higher. An schematic overview of the compartments is

given in Fig. 1. Each compartment was illuminated by six large windows (two off 1.25 m × 2.15 m and four off 1.25 m × 2.15). The compartments had one common wall and have both one wall adjacent to the exterior environment. The LED compartment was adjacent to the corridor and to a third compartment. The HPS compartment was adjacent to the corridor on two sides. Details of the surfaces are provided in Table 1. The floor of both compartments was covered with a plastic film, except in the first few metres adjacent to the corridor, where the floor consisted of concrete (see Fig. 1). Both compartments were equipped with thermal screens. The screens on the roof were Phormitex bright screens (Phormium, Lokeren, Belgium) and the screens at the wall were Obscura 10,070 R FR W screens (AB Ludvig Svensson, Kinna, Sweden).

Two separate heating systems were installed in each compartment. The major system was tube rail and consisted of a steel pipes with a diameter of 51 mm which was used for heating and fed by high-temperature water with an inlet temperature of up to 55 °C. There were two rail pipes per growing gutter with a length of 17.5 m. A second systems (designated growing pipes) had steel pipes with the same diameter but they were located between the plants and used for heating using low temperature water up to 40 °C. There were two growing pipes per gutter with a length of 17.5 m but one of these pipes can be lifted-up.

Both compartments were equipped with 180 μmol m^{-2} s^{-1} top-lights. To receive this intensity the LED compartment was equipped with 9 Signify DR/B LB GPL top-lights of 190 W each, 18 GreenPower DR/B LB LED top-lights of 190 W each and 33 GreenPower DR/W LB LED top-lights of 195 W (Signify, Eindhoven, the Netherlands). The HPS compartment was equipped with 16 GAN Electronic Gavita HPS lamps of 1035 W each. Thus, the overall power consumption of each system was 11,565 W for LED lamps and 16,560 W for HPS lamps. The lighting installations are installed 5.8 m above the floor. The climate in both compartments was controlled using both heating systems, opening and closing the windows and screens and influenced by the use of the lighting systems. Main parameters used by the climate control unit to realise an optimal climate were temperature and humidity.

Temperature, relative humidity and carbon dioxide (S-CO2) from both compartments and from the compartment adjacent to the LED compartment were monitored with a sensor present in the middle of each compartment. This logging device was positioned at the top of the plants, about 1.5 m above the ground. Temperature, wind direction, wind speed and

Table 1 – Surfaces of the greenhouse compartments.

(m^2)	LED compartment	HPS compartment
Common wall	120	120
Environment	208	208
Corridor	48	168
Third compartment	120	0
Ground	160	160

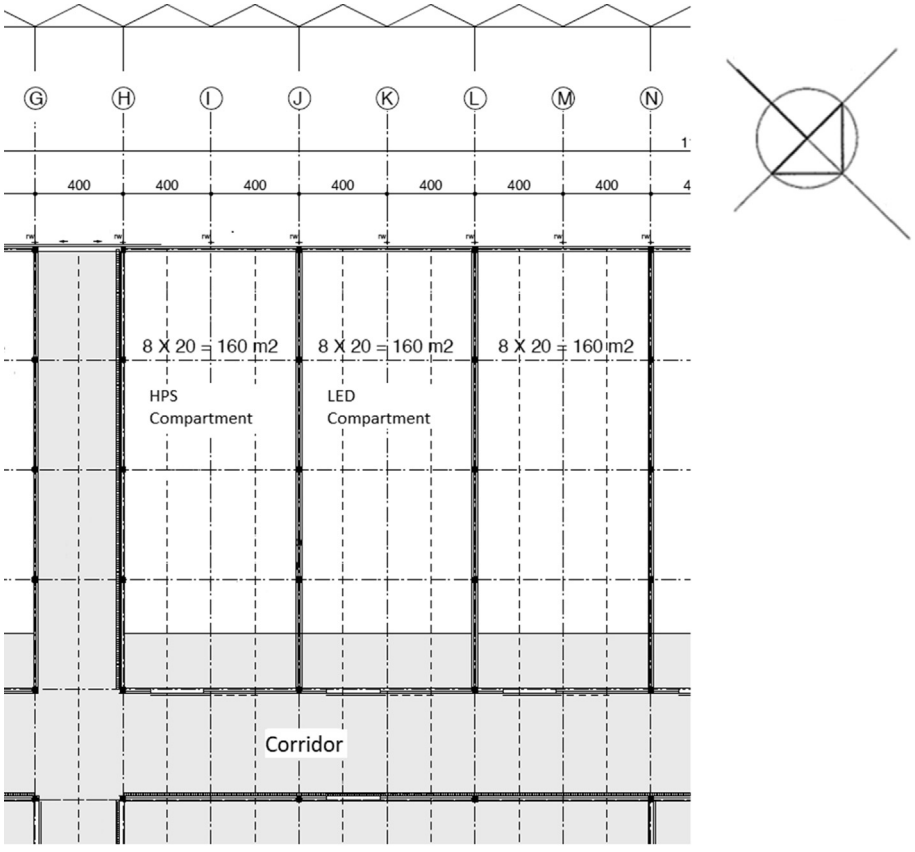

Fig. 1 – Cross section of the greenhouse. The grey colour indicates a concrete floor, while the white colour indicates a floor of ground covered with plastic sheeting.

incoming solar radiation is measured on the roof of the greenhouse (see outdoor sensors in Fig. 2). Relative humidity was measured about 1 m above the ground in a nearby field (see weather station in Fig. 2). The temperatures of the heating pipes were measured with thermocouples (MW80) as well as the flowrate in each pipe. The opening of the windows and screens was monitored as well. All data were measured at an interval of 60 s, using a data logging system (HortiMaX multima, Testo, Ternat, Belgium). Data were collected from November 1st until November 30^{t} 2018.

Cucumber plants (Cucumis sativus) were sown at 9 October 2018. The seedlings were transplanted on rockwool blocks (Grodan PlanTop Cube, Grodan, Roermond, The Netherlands). After 28 days, on 6 November 2018, the rockwool blocks were transmitted onto the rockwool mats (Grodan Master 0.1 m width x 1 m length, Grodan, Roermond, The Netherlands). Substrates were elevated 0.5 m above the ground with a drain gutter (width 0.31 m). A single-row system was used were plants are alternately guided to one side of the gutter. Distance between gutters is 1.6 m and the plant spacing was 250 mm, what resulted in a plant density of 2.5 plants per m^2 (one stem per plant). No extra stems were laid during the season. Plants were grown vertically in the greenhouse until they reached the support wires (1,5 m under the lamps). From then on, plants were moved clockwise around the gutter. Fruit thinning started alternately from the seventh fruit; one fruit per two nodes. Trellesing and pruning was carried out several times per week. The target yield was 7,5 fruits per week per m^2, based on 2,5 plants per m^2 x 6 leaves per week x alternate fruit thinning. Fruit was harvested form 5 December 2018 until 4 April 2019. Since plant growth was too generative at mid-December, fruit thinning was increased to 2:5 fruit:leaf ratio. There was a significant higher cumulative fruit yield in pieces (97 fruits m^{-2}) and mass (37 kg m^{-2}) in the HPS compartment in comparison to the LED compartment (88 fruits m^{-2} and 33.6 kg m^{-2}).

In the second experiment, vertical temperature data were measured in both compartments for two nights starting on May 15th, 2019. At this time, the LED compartment was empty, whilst cucumber plants were still present in the HPS compartment. Temperatures were monitored by a Testo data logger 177-T4 (Testo, Ternat, Belgium) equipped with RS PRO Type T Thermocouples. The sampling interval was 1 min. A vertical rope was used to position the temperature sensors with sensors placed at 500 mm above the lighting installation, at the same height as the installation, and at 500, 1600, 2500, 4000 and 5500 mm below the lighting installation. Experiments were performed at night to avoid the influence of sunlight on the temperature measurements. The lighting systems were turned on from 1 AM until 3 AM.

3. Integration scheme

The starting point is a generic scheme consisting of an estimator and model. The estimator is a least square estimator and the model is an ordinary differential equation. This scheme has two unknowns: the model parameters grouped in a vectorθ, and the model values T_t^{model} with index $t \in \{1, \ldots, N\}$. This scheme can be notated as

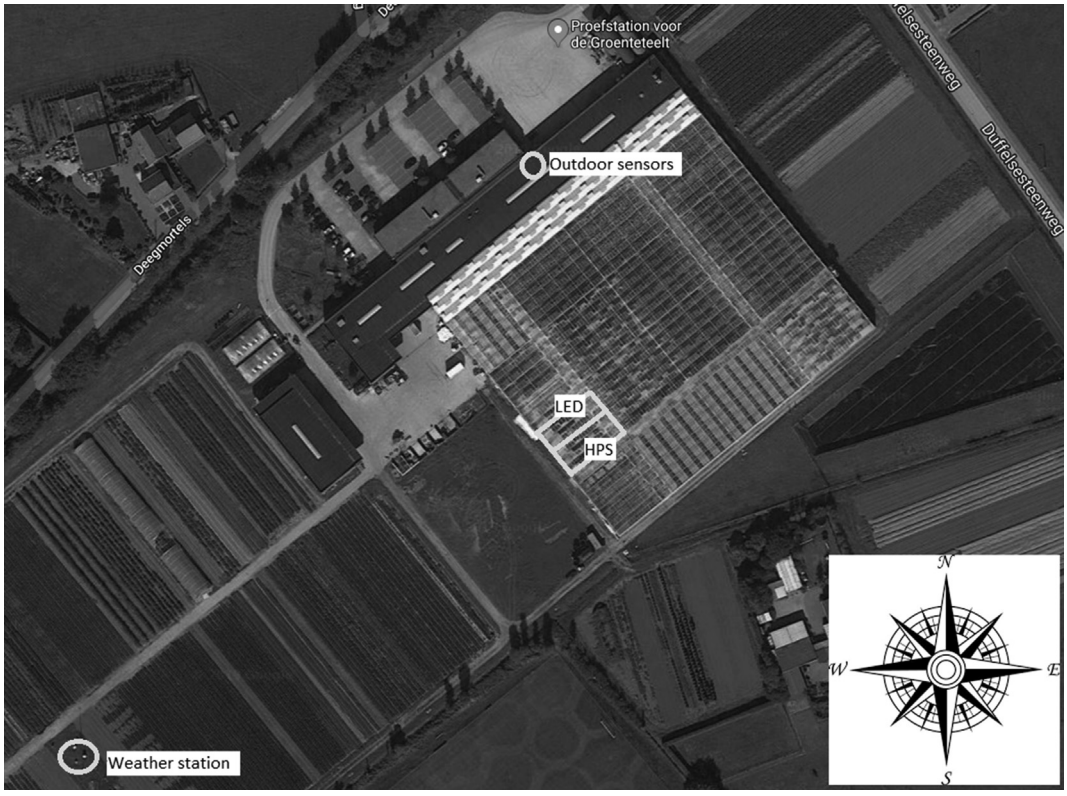

Fig. 2 – Greenhouse orientation with positioning of the LED and HPS compartment, the outdoor sensors and the weather station.

$$\left[\widehat{\theta}, \widehat{T}_t^{model}\right] = \arg\min_{\theta, T_t^{model}} \sum_{t=1}^{N} \left(T_t^{obs} - T_t^{model}\right)^2 \quad (1)$$

Subject to

$$\frac{dT^{model}}{dt} = f(T, \theta, \varepsilon^{obs}) \quad (2)$$

and possibly to other equality and inequality constraints. $\widehat{\theta}, \widehat{T}_t^{model}$ are the optimal parameters and model temperatures, T_t^{obs} the observed internal temperatures and ε^{obs} are a set of external observed variables, like outside temperatures. The function $f(T, \theta, \varepsilon^{obs})$ expresses the rate of change of temperature and depends on the model parameters, some external observed variables ε^{obs} and the state of the system T. The differential equation cannot usually be solved analytically, and a numerical approximation scheme must be used to estimate the values T_t^{model}. Therefore, a value for T and ε^{obs} has to be chosen in order to evaluate $f(T, \theta, \varepsilon^{obs})$. Depending on the integration method used, this value may change.

One of the most popular and simple integration schemes is the forward Euler integration scheme (Butcher & Goodwin, 2008)

$$\frac{T_{t+1}^{model} - T_t^{model}}{\Delta t} = f\left(T_t^{model}, \theta, \varepsilon_t^{obs}\right) \quad (3)$$

where the left-hand side is discretised and the right-hand side is evaluated in the last known state T_t^{model}. This equation can be rewritten to calculate the next state

$$T_{t+1}^{model} = T_t^{model} + f\left(T_t^{model}, \theta, \varepsilon_t^{obs}\right)\Delta t \quad (4)$$

In Eqs (1) and (2), θ and T_t^{model} are unknowns. The function $f(T_t^{model}, \theta, \varepsilon_t^{obs})$ often contains products of both. This makes it a nonlinear function in these unknowns. Such a nonlinear constraint requires the use of nonlinear solvers, including all problems with convergence and local minima.

However, this problem can be circumvented. It should be noted that interest is not specifically in the evolution of the temperature, since this evolution is already observed, but is in the parameters θ. Therefore, the rate of change function $f(T, \theta, \varepsilon^{obs})$ can be evaluated in T_t^{obs}. This leads to the following difference equation (still following the Euler integration scheme):

$$T_{t+1}^{model} = T_t^{model} + f\left(T_t^{obs}, \theta, \varepsilon_t^{obs}\right)\Delta t \quad (5)$$

The main advantage of this integration scheme is that the rate of change function $f(T_t^{obs}, \theta, \varepsilon_t^{obs})$ no longer contains products of unknowns. It is only a function of the unknown parameters θ. So, the three unknowns remaining in Eq. (5), i.e. T_t^{model}, T_{t+1}^{model} and θ, no longer appear as products. Therefore, difference equations such as Eq. (5) can be directly integrated into solvers for convex problems with affine constraints.

A disadvantage of this integration scheme is that additional measurement errors are introduced in the function $f(T_t^{obs}, \theta, \varepsilon_t^{obs})$ by means of both terms T_t^{obs} and ε_t^{obs}, instead of only by the term ε_t^{obs}. This lowers the precision of the estimator. To lower this impact, it is better to use the mid-point method (Butcher & Goodwin, 2008), instead of the explicit Euler integration. This leads to

$$T_{t+1}^{model} = T_t^{model} + f\left(T_{t+1/2}^{obs}, \theta, \varepsilon_{t+1/2}^{obs}\right)\Delta t \quad (6)$$

With

$$T_{t+1/2}^{obs} = \left(T_t^{obs} + T_{t+1}^{obs}\right)/2 \quad (7)$$

and

$$\varepsilon_{t+1/2}^{obs} = \left(\varepsilon_t^{obs} + \varepsilon_{t+1}^{obs}\right)/2 \quad (8)$$

By averaging over two subsequent observations, measurement error is reduced by a factor $\sqrt{2}$ (assuming normally independent identically distributed measurement noise).

It should be noted that the current version of this integration scheme cannot be applied in the following circumstances.

(i) If not all state variables are observed, the hidden states cannot be used to evaluate the right-hand side function.
(ii) It is assumed that the observations are sampled at the same sampling frequency as the time step in the difference equation. If the sampling frequency of the observations is lower than the model time step, observations should be interpolated. This might cause an additional interpolation error.

4. Model

This integration scheme was applied on a greenhouse model. Firstly, the different components of the model are described. Next, the estimator and solver are discussed and, finally, the method is expanded so that identical processes in multiple experiments can be characterised by the same parameters.

Conservation of energy, in the absence of ventilation, can be expressed as

$$c\frac{dT^{model}}{dt} = \sum_{i=1}^{M} Q^i \quad (9)$$

where c is a heat capacity and the right-hand side summates over all heat sources and sinks. This differential equation can be discretised with respect to time

$$T_{t+1}^{model} = T_t^{model} + \frac{1}{c}\sum_{i=1}^{M} Q_{t+1/2}^{i}\Delta t \quad (10)$$

This equation shows how temperature will evolve over a time step Δt from a given temperature T_t^{model}. A series of heat sources and sinks $Q_{t+1/2}^{i,obs}$ with $i \in \{1, \ldots, M\}$ are considered and discussed below.

4.1. Ventilation

Ventilation is now incorporated in this difference equation, based on the work of van Henten, Bontsema, Kornet, and Hemming (2006) and de Brauwere et al. (2007): if a fraction of

interior air α_t, is replaced by environmental air, then the temperature changes to

$$T_{t+1}^{obs} = (1-\alpha_t)T_t^{obs} + \alpha_t T_t^{exterior} \quad (11)$$

By measuring, the interior and exterior states, one can easily derive the fractions α_t. Evidently, this fraction is restricted to

$$0 \le \alpha_t \le 1 \quad (12)$$

If heat sources and sinks are included, the temperature evolution in the next time step is given by combining Eqs. (10) and (11):

$$T_{t+1}^{model} = (1-\delta_t^{vent})T_t^{model} + \delta_t^{vent}\left((1-\alpha_t)T_t^{obs} + \alpha_t T_t^{exterior}\right) + \frac{1}{mc}\sum_{i=1}^{M} Q_t^i \Delta t \quad (13)$$

With δ_t^{vent} a binary variable (zero if windows are closed and one if windows are open).

4.2. Conduction losses to the environment

Heat losses to the environment are assumed to be proportional to the temperature difference between the greenhouse and the environment (van Henten, 1994). Many greenhouses are equipped with screens to lower heat losses during night. The position of these screens is modelled by a binary parameter δ_t^{screen}, which is assumed to be known and can take two values, namely 1: screen closed and 0: screen open.

$$Q_t^1 = \left(\theta_1 \delta_t^{screen} + \theta_2(1-\delta_t^{screen})\right)\left(T_t^{obs} - T_t^{exterior}\right)S^{exterior} \quad (14)$$

Q_t^1 is the heat exchanged with the environment at time t, θ_1 (heat losses when screens are closed) and θ_2 (heat losses when screens are open) are two unknown parameters, which are assumed to be constant in time, T_t^{obs} and $T_t^{exterior}$ are the observed interior and exterior temperatures, respectively, and $S^{exterior}$ is the surface that connects the greenhouse to the open-air.

4.3. Transmission losses to neighbouring compartments

Heat losses to neighbouring compartments are modelled similarly. Usually no screens are present between neighbouring compartments, so

$$Q_t^2 = \sum_{j=1}^{J} \theta_3 \left(T_t^{obs} - T_t^{neighbor\ j}\right) S^{neighbor\ j} \quad (15)$$

where Q_t^2 is the heat exchanged with all neighbours at time t. $T_t^{neighbor\ j}$ is the temperature at neighbour j at time t, with $j \in \{1,\ldots,J\}$ and J the number of neighbours, $S^{neighbor\ j}$ is the common surface between the greenhouse compartment and its neighbour j. For simplicity, it is assumed that the physical properties of the walls that separate the greenhouse from its neighbours are all identical. If this would not be the case, the parameter θ_3 should become neighbour dependent.

4.4. Latent heat

Condensation and evaporation in greenhouses have been extensively studied (e.g. Pieters and Deltour (1997a) or Pieters and Deltour (1997b)). In these studies, a model was used to simulate the effect of condensation and evaporation on the auxiliary heating requirements, on the inside air humidity and temperature and on the vegetation temperature. Such models are impractical for this study, since many states, such as the soil and roof temperature are usually not observed. Therefore, heat consumption by evaporation and condensation can be estimated from changes in the absolute humidity (g m^{-3}). A distinction is made between the two situations:

(i) When vents are closed, all changes in humidity must be related to changes in latent heat and influence the heat present in the system.
(ii) If vents are open, two processes are taken into account
   a. Changes in internal humidity
   b. Humidity exchange with the environment

$$Q_t^3 = (1-\delta_t^{vent})Q_t^{internal} + \delta_t^{vent}\left((1-\alpha_t)Q_t^{internal} + \alpha_t Q_t^{external}\right) \quad (16)$$

With

$$Q_t^{internal} = -\frac{LV}{\Delta t}\left(\chi_{t+1/2}^{obs} - \chi_{t-1/2}^{obs}\right) \quad (17)$$

$$Q_t^{external} = -\frac{LV}{\Delta t}\left(\chi_{t+1/2}^{obs} - \chi_{t-1/2}^{exterior}\right) \quad (18)$$

And L the latent heat, V the volume of the greenhouse and Δt the sample period. $Q_t^{internal}$ is the positive or negative contribution of changes in humidity inside the greenhouse compartment, to the internal heat. Note that Eqs. (17) and (18) are approximations: the humidity $\chi_{t+1/2}^{obs}$ is a result of internal evaporation and condensation and of humidity exchange with the environment.

4.5. Solar heating

The heat provided by the sun was given by

$$Q_t^4 = \theta_4 P_t^{Sun} S^{ground} \quad (19)$$

With $\theta_4$unknown parameters that account for the fraction of heat from the sun that is captured by the greenhouse compartment. The remaining fraction is scattered or reflected on the roof or is used to heat neighbouring compartments, P_t^{Sun} represents the intensity of the sun (W m^{-2}).

4.6. Heating system

The model used for the heating system depends on the available data. If the necessary data are available, the amount of heat provided by the heating systems can be calculated, e.g. with

$$Q_t^5 = \theta_5 \dot{m}_t c_p \left(T_t^{in} - T_t^{out}\right) \tag{20}$$

with T_t^{in} the input temperature of the heating system, T_t^{out} the output temperature.

Often greenhouses are equipped with a high and low temperature heating system. If several heating systems are present, each heating system should be modelled separately and described by its unique parameter(s). Obviously, the parameter θ_5 should be equal to one.

4.7. Lighting system

The lighting system was modelled as

$$Q_t^6 = \theta_6 P_t^{light} \tag{21}$$

with P_t^{light} the electric power provided to the lighting system. The unknown parameter θ_7 tells which fraction of the electric power ends up as heat in the greenhouse. If multiple lighting systems are present, each should be modelled separately and will be described by its own parameter(s). If shading screens are present, these could be incorporated similarly to the thermal screens.

4.8. Parameter estimator

As is pointed out above, estimating the unknown parameters was carried out by casting this problem as a convex problem with affine constraints. The problem can be formulated as follows

$$\left[\hat{\theta}, \hat{T}_t^{model}, \hat{\alpha}_t\right] = \arg\min_{\theta, T_t^{model}, \alpha_t} \sum_{t=1}^{N} \left(T_t^{obs} - T_t^{model}\right)^2 \tag{22}$$

Subject to Eqs. (12–21).

This least squares estimator is subject to a set of affine equality and inequality constraints. Therefore, it is a convex problem that can be solved very efficiently using e.g. CVX, a package for specifying and solving convex programs (Grant, Boyd, & Ye, 2008; Grant & Boyd, 2008).

4.9. Estimation of the mass and heat capacity product

A problem with this model is that the heat capacity c is not known. If the term c is added to the parameters to be estimated, the problem is no longer solvable. To circumvent this problem, the ratio θ/c can be estimated. Since the parameter associated with the heating system θ_5 should be one, the heat capacity can be identified from the estimate of θ_5.

$$c^{update} = c^{guess} / \theta_5 \tag{23}$$

This equation can be used to find correct values for the overall mass and heat capacity of the greenhouse.

4.10. Identical parameters in multiple experiments

If the aim of an experiment is to compare two specific processes, while other processes remain identical, it is convenient to combine these multiple experiments. This can for example be done by repeating an experiment in a greenhouse or by using two compartments in one greenhouse. The parameters of the identical processes should remain the same. In this paragraph, the procedure is expanded so that a set of parameters can be shared between estimators. Assuming that a subset of the parameters is identical, $\Omega \subseteq \{1,\ldots,K\}$, with $K = 6$ in this case. This subset is indexed with $k \epsilon \Omega$. This can be formulated as

$$\theta_k^m = \overline{\theta}_k \tag{24}$$

With θ_k^m the parameter k in compartment m and $\overline{\theta}_k$ the mean value of parameter k over all compartments. If this condition is fulfilled, all parameters are equal to their mean and thus to each other. This common constraint can be incorporated in the objective function of estimator m, which becomes

$$\left[\hat{\theta}_i^m, \hat{T}_t^{model}, \hat{\alpha}_t\right] = \arg\min_{\theta_i^m, T_t^{model}, \alpha_t} \sum_{t=1}^{N} \left(T_t^{obs} - T_t^{model}\right)^2 + \sum_{k \epsilon \Omega} \lambda_k^m \left(\theta_k^m - \overline{\theta}_k\right) \tag{25}$$

Subject to Eqs. (12–21).

If constraint Eq. (24) is fulfilled, the latter term vanishes. This objective function is called the Lagrangian and λ_k is the Lagrange multiplier. This Lagrange multiplier cannot be chosen arbitrarily but has a unique value. To find this, the Lagrange multipliers can be initialised as zeros. In this case, the unconnected cases is solved for each compartment. Next, the mean parameter value can be calculated, and the Lagrange multipliers can be updated as follows

$$\lambda_k^{m,l+1} = \lambda_k^{m,l} + \rho\left(\theta_k^m - \overline{\theta}_k\right) \tag{26}$$

With ρ a small number and l the iteration number. This scheme can be iterated until λ_k^m has converged. However, in practice, this may take a while. Therefore, the Lagrangian is often augmented with a quadratic term, which speeds up convergence and makes the numerical scheme more robust (Boyd, 2011). The augmented Lagrangian is given by

$$\left[\hat{\theta}_i^m, \hat{T}_t^{model}, \hat{\alpha}_t\right] = \arg\min_{\theta_i^m, T_t^{model}, \alpha_t} \sum_{t=1}^{N} \left(T_t^{obs} - T_t^{model}\right)^2 + \sum_{k \epsilon \Omega} \lambda_k^m \left(\theta_k^m - \overline{\theta}_k\right) + \frac{\rho}{2} \sum_{k \epsilon \Omega} \left(\theta_k^m - \overline{\theta}_k\right)^2 \tag{27}$$

subject to Eqs. (12–21).

This quadratic term also vanishes when the coupling constraint is met (Eq. 24). This optimisation scheme is based on the alternating direction method of multipliers (ADMM (Boyd, 2011)), more particularly on the so-called consensus algorithm, because all estimators must find a consensus on the common parameters. In practice, convergence is reached when the largest difference between the common parameters is less than 1%.

Note that estimating the heat capacity c with the procedure described above can only be combined with this method if the heat capacity in all compartments is identical. If they differ, the ratios θ/c will differ for each compartment and no consensus can be found.

4.11. Evaluation of the model on a test set

Applying the method to a test set, the model needs to be evaluated while ventilation variables need to be identified. At the same time, the model parameters θ, should remain the same. This can be achieved by applying the method on the novel data with one additional constraint

$$\left[\hat{T}_t^{model}, \hat{\alpha}_t\right] = \arg \min_{T_t^{model}, \alpha_t} \sum_{t=1}^{N} \left(T_t^{obs} - T_t^{model}\right)^2 \quad (28)$$

Subject to Eqs. (12–21) and to $\theta = \hat{\theta}$, where $\hat{\theta}$ are the previously estimated parameters. The two additional terms in the augmented Lagrangian are fulfilled, so they do not contribute to the cost function and can be ignored.

5. Procedures

The observations were split into two sets. The first set was called the training set and was used to evaluate the proposed model and estimate model parameters. The second set was called the test set and was used to evaluate the final performance of the model. The test set was not used during the evaluation and tuning of the model. The training set comprised all data from November 9th, 2018 at 08:00 to November 16th at 06:41. The test set consisted of data from November 16th at 15:00 to November 23rd at 13:40. All data were averaged so that the sample period was always 20 mins.

Since the material of the walls was identical in both compartments, it was assumed that the thermal resistance of all external walls and roof and the thermal resistance of internal walls were identical. Therefore, these parameters were coupled in the estimator by a Lagrange multiplicator. Other parameters, such as the influence of sun light and the lighting system may differ in each compartment.

6. Results

6.1. Case study

6.1.1. The training set and estimated parameters

Measurements, model and ventilation rates are shown in Figs. 3 and 4. Temperatures vary between 18 and 26 °C, with higher temperatures during the day and colder periods during night. The most striking variation was the sudden temperature drops each evening. The model was well able to follow the temperature variations. As a measure of the quality of the model, the ratio between the root-mean-square-error (RMSE) to the root-mean-square of the observations was used. This is a measure for that variation that cannot be described by the model. Values were 1.8 and 2.0% for respectively the LED and HPS compartments. Overall, the model was able to explain more than 95% of the variation in temperature.

The ventilation rate was very similar in both compartments. During the night windows were closed and the ventilation rate was, evidently, absent. During most of the day ventilation varied between 0 and 10%. Most evenings, a spike in ventilation was observed, which coincide with the sharp temperature drop. These spikes varied between 75 and 300 W m^{-2}.

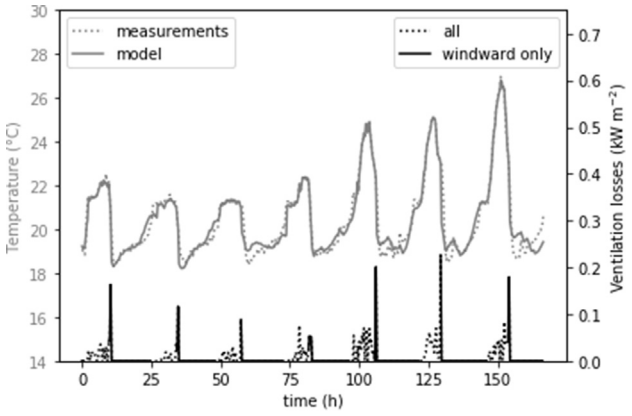

Fig. 3 – Measurements and model predictions for the training set of the LED compartment. The grey lines are respectively the measured and modelled indoor temperature and the black lines represents the estimate ventilation losses. Dotted black lines are ventilation losses when the leeward window was opened (and possibly also the windward window), while the full black line represents moments when only the windward window was opened.

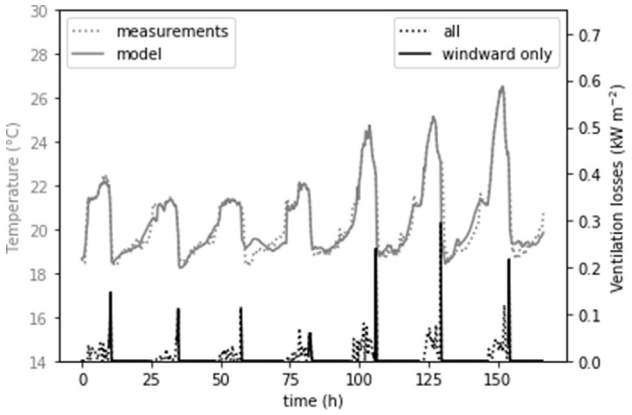

Fig. 4 – Measurements and model predictions for the training set of the HPS compartment. The grey lines are respectively the measured and modelled indoor temperature and the black lines represents the estimate ventilation losses. Dotted black lines are ventilation losses when the leeward window was opened (and possibly also the windward window), while the full black line represents moments when only the windward window was opened.

The estimated parameters are given in Table 2. The heat capacity of the greenhouse was estimated to be 0.083 kWh m^{-2} K^{-1}. The overall heat transfer coefficient doubled when the thermal screens were in use (parameter decrease from 6.1 to 2.7 W m^{-2} K^{-1}). Heat exchange with neighbouring compartments, or the corridor, was characterised by a conduction parameter of 2.0 W m^{-2} K^{-1}. For the LED and HPS compartment, respectively 76 and 82% of the sunlight ends up in the greenhouse as heat.

For the lighting installations, 37% of the electric energy for the LED system and 24% for the HPS system eventually ends as heat in the compartments. These results indicate that LEDs

Table 2 – Estimated parameters.

		Parameter	LED compartment	HPS compartment	Values found in Literature
Heat capacity (kWh m^{-2} K^{-1})		c	0.083		
Overall heat transfer coefficient (W m^{-2} K^{-1})	Screen closed	θ_1	2.7		2.94 (Geelen et al., 2018)
	Screen open	θ_2	6.1		7 (Geelen et al., 2018)
Heat exchange with neighbours (W m^{-2} K^{-1})		θ_3	2.0		–
Heat absorption of solar radiation (%)		θ_4	76	82	66 (Pieters, 1999), 70–80 (Geelen et al., 2018), 74 (Stanghellini et al., 2019)
Heating system parameter (%)		θ_5	100	100	–
Heat absorption of artificial light (%)		θ_6	37	24	–

are more efficient at transforming electric power into heat than HPS installations at the location of the measurement box.

6.1.2. *Description of the test set*

The observed temperatures, the modelled temperatures and the ventilation rates of the test set are shown in Figs. 5 and 6. Temperature ranges were similar to the training set and varied between 18 and 26 °C. The model was able to follow temperature variations, but the model error was obviously more severe than in the training set. During most days observations followed predictions well, but during the third and fourth days (hours 48 to 96) some severe deviations were observed. The match was in generally better during day time than at night. This is probably because the estimation of ventilation rates allowed the model to follow observations more closely. Overall, the model still explained over 90% of the variability seen in the observations. When the test dataset was being collected, the compartments were less well ventilated, but the estimated ventilation in both compartments was still similar (the correlation coefficient between both ventilation rates was still 90%). Within each compartment the estimated ventilation rates were similar to those estimated in the training set: no windows were opened at night, during day time ventilation rates varied between 0 and 100 kW m^{-2} with a peak up to 200 kW m^{-2} before nightfall.

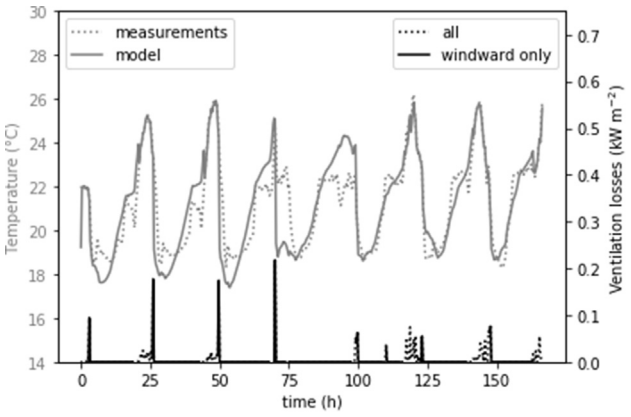

Fig. 5 – Measurements and model predictions for the test set of the LED compartment. The grey lines are respectively the measured and modelled indoor temperature and the black lines represent the estimate ventilation losses. Dotted black lines are ventilation losses when the leeward window was opened (and possibly also the windward window), while the full black line represents moments when only the windward window was opened.

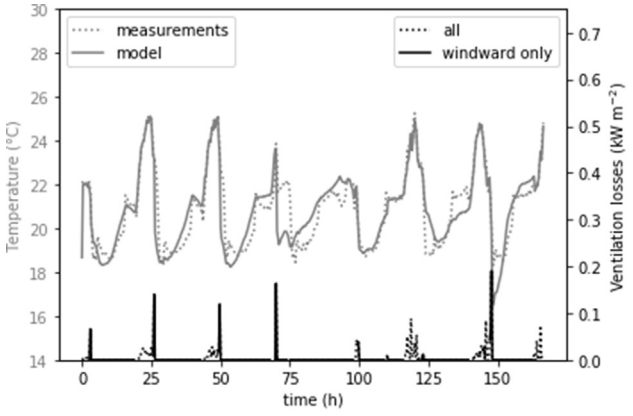

Fig. 6 – Measurements and model predictions for the test set of the HPS compartment. The grey lines are respectively the measured and modelled indoor temperature and the black lines represent the estimate ventilation losses. Dotted black lines are ventilation losses when the leeward window was opened (and possibly also the windward window), while the full black line represents moments when only the windward window was opened.

6.1.3. *Estimated heat flows*

Table 3 shows the daily averaged heat flows (W m^{-2}). Heat delivered by the heating installation was the main source of heat. It was 72.9 W m^{-2} and was 7.5% greater for the LED compartment. The difference in heat absorption of sun light was 2.5 W m^{-2} or about 8.7% smaller for LED. The overall amount of heat provided by the lighting systems was similar in both compartments, although the LED compartment provided almost 1 W m^{-2} more. The main heat sink was losses to the environment. Since both compartments had very similar temperatures, identical thermal resistance parameters and identical surfaces, the heat losses were almost identical. The small difference of 1.7 W m^{-2} is due to the slightly lower temperature of HPS compartment. This average temperature difference was 0.3 °C. Heat losses to other compartments and the corridor differs quite a lot: 2.2 W m^{-2} or over 27%.

Table 3 – Overview of the average heat flows.

($W\ m^{-2}$)	LED compartment	HPS compartment	Relative difference w.r.t. LEDs (%)
Heating system	72.9	67.4	7.5
Sun	28.6	31.1	8.7
Lighting	14.8	13.9	6.1
Environment	−85.6	−83.9	2.0
Neighbours	−8.0	−10.2	27.5
Evaporation	−1.0	−1.2	20
Condensation	0.67	0.91	35
Humidity exchange	−0.2	−0.2	0
Ventilation	−25.6	−21.1	17.6
Remaining heat	3.5	3.3	5.7

It should be noted that crop growth was less tenuous and better, and the leaves were bigger in HPS compartment, which can explain the difference in latent heat. Ventilation losses were greater for the LED compartment. Not only was the total amount of losses large, but also the difference in losses between both compartments was large. Evaporation, condensation and humidity exchange during ventilation were all small in this experiment, which is not surprising since the plants were all still small. Overall heat sources and sinks must be in equilibrium. Therefore, the remaining heat is the amount of heat needed to evolve from the initial to the final temperature.

6.2. *Vertical temperature measurements*

It is well known that LED lamps are more efficient than HPS lamps. However, the parameters estimated here seem to show the contrary. In Table 2 LEDs convert 37% to heat, whilst the HPS system converted only 24%, while both produce the same amount of PAR light. To understand this difference, a second experiment was performed. Here, both compartments were equipped with vertical temperature sensors. During night, both lighting systems were switched on and the changes in temperature distribution were measured. The data of this experiment is shown in Fig. 7. The LED compartment was on the average 1 °C warmer. Therefore, the figure shows the temperature changes relative to midnight. Once both lighting systems were switched on, temperature started to increase in all air layers. However, the steady state temperature under lighting was very different for both systems. At 4 m below the lighting, where the climate box and crops were, both lighting systems contributed the same to the temperature (+1 °C). Higher p, closer to the lamps, the HPS lamps heat up the air more than the LEDs. This difference was more pronounced at the height of the lighting installation. Half a meter below the lighting installations, the LED heats the air still 1 °C, while the HPS installation increased the temperature by 3 °C. Once the lighting installations were switched off the temperatures decreased to an equilibrium temperature (see Fig. 7).

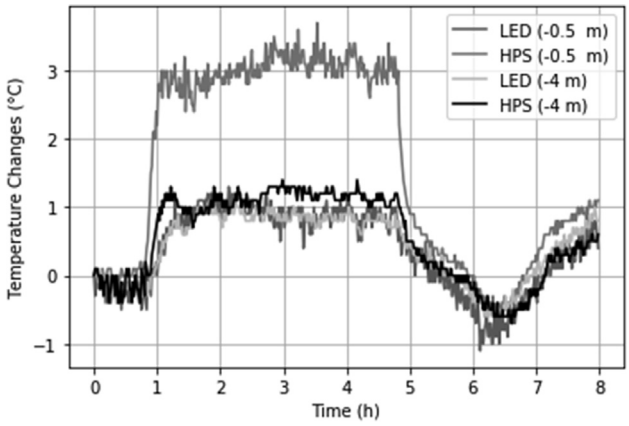

Fig. 7 – Temperature trends under LED and HPS lamps at night. The position is relative to the lighting installation (−0.5 m is half a metre below the lighting installation, while −4 m is about 1.5 m above the ground, where the temperature sensor of the climate box was placed.

7. Discussion

Firstly, the ventilation and soil interaction processes are discussed. This is followed by the most important model parameters which are compared with literature.

7.1. *Ventilation*

7.1.1. *Noise and model errors*

The ventilation rate estimated in this paper was not parameterised, so for every additional observation a new ventilation variable was estimated. In other words, the number of variables to be estimated increased with the number of measurements when ventilation was present. Therefore, measurement noise could not be averaged out, as is usually done in parameter estimation techniques. Consequently, most measurement noise, and possibly model errors, were propagated into the ventilation rate. It seems, however, that the model matched the observations quite well, even when no ventilation was present. So, we might expect that this noise and error propagation is reasonably small.

7.1.2. *Peaks in the ventilation rate*

One of the most striking features of the estimated ventilation were the large peaks that occurred every evening around sunset. These ventilation peaks were caused by the climate control settings. During the day, ventilation was usually caused by opening the leeward windows. However, before sunset, the leeward windows were closed and the windward

windows were opened (shown by the full and dotted lines in the ventilation rates in Figs. 3–6). This caused a large ventilation peak, resulting in a replacement of humid internal air by relatively dry external air. These peaks in ventilation were therefore deliberately caused by the climate controller.

7.1.3. Interpretation of the ventilation rate

It is difficult to come up with a test to judge the estimated ventilation rates. De Jong and Bot (1992) proposed a model for ventilation based on the wind speed and window opening: for a given window opening, the ventilation rate was proportional to the wind speed and for small window openings the ventilation rate was proportional to the window opening. For larger window openings, the ventilation rate saturated. Parra, Pérez, Baeza and Motero (2004) added a buoyancy term to ventilation that was proportional to the temperature difference between inside and outside the greenhouse. Of course, the dynamics of ventilation will differ, particularly if windows are opened on the windward or leeward side or in both directions.

Wind direction and orientation with respect to the windows was also taken into account. Hence, to test the accuracy of the ventilation rate, five features were considered:

- Product of windspeed and window opening
- Windspeed component parallel to the window opening
- Windspeed component perpendicular to the window opening
- Wind direction
- Temperature difference between inside and outside the greenhouse.

A multi-linear regression model was used to map the estimated ventilation rate on these features and this for the three types of ventilation. Results are shown in Table 4. If both window types were opened, or if only the leeward windows were opened, the ventilation rate was reasonably well described by these basic features. When only the windward windows were opened, no proportional relationship was identified between estimated ventilation rates and basic features. This does not necessarily mean that the model failed in this situation; it just means that the relationship was more complex than can be discovered by a multilinear regression model.

7.2. Soil interaction

Several models were tested to identify heat exchange with the soil. In the first model, no heat is exchanged at all, i.e. $Q_t^7 = 0$. In the second model, heat exchange is constant, i.e. $Q_t^7 = \theta_7$. In the third model, heat exchange is proportional to the temperature difference between the compartment and the soil, i.e. $Q_t^7 = \theta_7(T_t^{obs} - T^{soil})$, with $T^{soil} = 20\ ^\circ C$. In a fourth model, the heat exchange is assumed to be period $Q_t^7 = \theta_9 \sin(\omega t + \theta_{10})$, with $\omega = 2\pi/(24h)$. In the latter model, it was assumed that the periodic heat exchange has a daily period and that the average heat exchange over a day is zero. The shape of the heat exchange curve is therefore assumed to be sinusoidal with unknown amplitude and unknown phase. It appeared that none of the parameters estimated in any of these models resulted in a significant heat exchange with the soil. For that reason, heat exchange with the soil was not considered in this case study.

Table 4 – Correlation coefficients of multi-linear regression model.

	LED compartment	HPS compartment
Both windows are opened	92%	75%
Leeward windows are opened	58%	59%
Windward windows are opened	44%	45%

7.3. Parameters

7.3.1. Heat capacity

The heat capacity c was estimated at 0.083 kWh m^{-2} K^{-1}. This is about 42 times the heat capacity of the air in the compartment. This means that the estimators did not only take the air into account, as observed by the climate box, but also the plants, tables, construction, part of the soil, etc. This heat capacity also influences the modelled ratio between latent and sensible heat. Under these circumstances sensible heat is the much more important.

7.3.2. Overall heat transfer coefficient

The estimated overall heat transfer coefficient, for the greenhouse without and with screens was 6.1 and 2.7 W m^{-2} K^{-1} respectively. Geelen et al. (2018) noted that "the energy transfer through a greenhouse cover is always an combined effect of conductance through the material itself and convection depending on air temperatures and air movement on either side of the material". For an accurate calculation of the energy flow, all these factors are needed. However, in practice it is sufficient to use an effective U-value that applies under "normal conditions". A typical value is about 7 W m^{-2} K^{-1} (Geelen et al., 2018). The value of $U^{greenhouse}$ =6.1 W m^{-2} K^{-1} is close to this value, which confirms that the model applied here matches reasonably well. The effective U-value of the screen can be computed from the combined effective U-value of 2.7 W m^{-2} K^{-1} (i.e.$1/U^{combined} = 1/U^{greenhouse} + 1/U^{screen}$). It's value is 5.1 W m^{-2} K^{-1}, which is again in close agreement with a 6 W m^{-2} K^{-1} mentioned by Geelen et al. (2018).

7.3.3. Solar radiation

Two distinct parameters were found for the efficiency of the greenhouse to capture solar radiation. The LED compartment had an efficiency of 76%, while for the HPS compartment had an efficiency of 82%. Pieters (1999) studied the solar energy input in greenhouses. An interesting rule of thumb he found is that two third of the solar energy is captured by the greenhouse. Geelen et al. (2018) estimated this parameter to be between 70 and 80% and Stanghellini et al. (2019) estimated this parameter to be 74%. The parameters estimated here agree with these findings. An explanation is that both compartments do not experience the same shadow from neighbouring compartments. This may explain the difference in efficiency with respect to solar radiation.

7.3.4. *Lighting installation*

One of the most remarkable results found here is that the thermal efficiency used by the model is higher for LEDs (37%) then for HPS lamps (24%), while it generally known that LED lamps are more efficient than HPS lamps. Firstly, it has to be noted that this parameter relates the electric power of the lighting installations with changes in temperature measured at the climate box. Secondly, both lighting installation were selected to provide the same amount of PAR light (180 $\mu mol\ m^{-2}\ s^{-1}$). Therefore, a 11,565 W LED installation and a 16,512 W HPS installation were used. By applying this efficiency parameter, these installations produced respectively 4279 and 3962 W heat. Both values are of the same order of magnitude. To substantiate this result, the temperature was measured at different positions under the lights. These measurements confirmed that both lighting installations provided the same amount of heat at a height of 1.5 m above the ground and therefore that these efficiencies appear accurate. However, the results show more. The HPS installation heats up the top layers in the greenhouse much more than the LED installation. So, in fact there is no contradiction. LED lamps are overall more efficient, but the additional heat that is produced by HPS lamps does not reach the climate box, nor the plants, but is concentrated in the top air layers of the greenhouse. A similar effect was noticed by Geelen et al. (2018): "Thus, supplemental light in the form of LED top lights results in a completely different temperature profile compared to HPS lighting, because the latter brings much more heat to the top of the plants." When the windows were opened, this warm air layer left the greenhouse. This model cannot make a distinction between energy leaving the compartment directly as light or by this pathway.

In the current model, only one state is used to describe the greenhouse and that is the temperature measured by the climate box. As Fig. 7 shows, the temperature under the roof can be several degrees warmer than at the climate box. In the compartments studied here, this did not have a large impact on the temperatures observed by the plants or climate box, but if, for example, fans are used, these higher temperatures could reach the plants. In such situations, it may be useful in modelling this scenario to incorporate multiple states. A consequence is that a single temperature measurement, without vertical variations may mean that this single state is not always representative of the complete compartment. Firstly, plants do not experience the higher temperatures in the top layers. And secondly, a vast majority of the greenhouse is not equipped with vertical temperature sensors. Therefore, the application range of a data driven model based on vertical temperatures would be much smaller.

8. Conclusion

A method was proposed to make parameters estimates in grey-box models convex. This makes the estimator procedure independent of initial values and robust against local minima in the optimisation procedure. The method was expanded so that identical processes in different experiments can be characterised by the same parameters.

This method was applied on experimental data from a greenhouse, where two compartments were compared. All circumstances were kept similar, expect the lighting installations. One compartment was equipped with HPS lamps, while the other was equipped with LED lamps.

A particular complication with greenhouse models is ventilation: a significant amount of heat is lost through ventilation. Ventilation losses were not measured explicitly, so this heat sink had to be estimated from temperature changes too. This process was incorporated well.

Processes that were included conduction losses to the environment and to neighbouring compartments, latent heat absorption due to transpiration, heating by the sun, heating by the heating system(s) and also by the lighting systems. Parameters found for these common processes were close to values found in literature. No significant heat losses to the ground could be identified.

One of the noteworthy outcomes of this experiment was that the amount of heat produced by both lighting systems, as is observed by the climate box and plants was identical. To confirm this result, the vertical temperature distribution under both lighting systems were measured. It appeared that the HPS system heated up the upper air layers more than the LED system. However, a few metres below the lighting installations both systems provided the same amount of heat.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was carried out in the research project GLITCH. GLITCH is supported by Interreg Flanders-Netherlands V – European Fund for Regional Development, European Union, the cross-border cooperation program with financial support from the European Fund for Regional Development. Additionally, the project is supported by the Agency for Innovation and Entrepreneurship (VLAIO), Belgium, the Province of Antwerp, Belgium, the Flemish Cabinet for Environment, Nature and Agriculture, Belgium, the Province of Limburgthe Netherlands and the Dutch Ministry of Economic Affairs, the Netherlands.

REFERENCES

Akman, D., Akman, O., & Schaefer, E. (2018). Parameter estimation in ordinary differential equations modeling via particle swarm optimization. *Journal of Applied Mathematics, 2018*, 1–9. https://doi.org/10.1155/2018/9160793.

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., & Diehl, M. (2018). CasADi: A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation, 11*(1), 1–36. https://doi.org/10.1007/s12532-018-0139-4.

Bock, H. G., & Plitt, K. J. (2017). A multiple shooting algorithm for direct solution of optimal control problems *. *IFAC Proceedings Volumes, 17*(2), 1603–1608. https://doi.org/10.1016/s1474-6670(17)61205-9.

Boulard, T., Haxaire, R., Lamrani, M. A., Roy, J. C., & Jaffrin, A. (1999). Characterization and modelling of the air fluxes induced by natural ventilation in a greenhouse. *Journal of Agricultural and Engineering Research, 74*(2), 135–144. https://doi.org/10.1006/jaer.1999.0442.

Boyd, S. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning, 3*(1), 1–122. https://doi.org/10.1561/2200000016.

de Brauwere, A., Jacquet, S. H. M., De Ridder, F., Dehairs, F., Pintelon, R., Schoukens, J., et al. (2007). Water mass distributions in the Southern Ocean derived from a parametric analysis of mixing water masses. *Journal of Geophysical Research Oceans, 112*(2), 1–16. https://doi.org/10.1029/2006JC003742.

Butcher, J., & Goodwin, N. (2008). *Numerical methods for ordinary differential equations*.

De Jong, T., & Bot, G. P. A. (1992). Air exchange caused by wind effects through (window) openings distributed evenly on a quasi-infinite surface. *Energy and Buildings, 19*(2), 93–103.

Ding, A. A., & Wu, H. (2013). Estimation of ordinary differential equation parameters using constrained local polynomial regression. *Statistica Sinica, 24*, 1613–1631. https://doi.org/10.5705/ss.2012.304.

Frausto, H. U., Pieters, J. G., & Deltour, J. M. (2003). Modelling greenhouse temperature by means of auto regressive models. *Biosystems Engineering, 84*, 147–157. https://doi.org/10.1016/S1537-5110(02)00239-8.

Geelen, Voogt, & van Weel. (2018). *Plant empowerment the basic principles*. Letsgrow.com.

Grant, M. C., & Boyd, S. P. (2008). Graph implementations for nonsmooth convex programs. In *Recent advances in learning and control* (pp. 95–110). Springer.

Grant, M., Boyd, S., & Ye, Y. (2008). *CVX: Matlab software for disciplined convex programming*.

Guzmán-Cruz, R. (2013). Modelling greenhouse air temperature using evolutionary algorithms in auto regressive models. *African Journal of Agricultural Research, 8*(3), 251–261. https://doi.org/10.5897/AJAR12.280.

He, F., & Ma, C. (2010). Modeling greenhouse air humidity by means of artificial neural network and principal component analysis. *Computers and Electronics in Agriculture, 71*(Suppl. 1). https://doi.org/10.1016/j.compag.2009.07.011.

van Henten, E. J., Bontsema, J., Kornet, J. G., & Hemming, J. (2006). *On-line schatting van het ventilatievoud van kassen* (Vol. 62). Retrieved from http://edepot.wur.nl/38612.

Kimball, B. A. (1973). Simulation of the energy balance of a greenhouse. *Agricultural Meteorology, 11*, 243–260.

Mashonjowa, E., Ronsse, F., Milford, J. R., & Pieters, J. G. (2013). Modelling the thermal performance of a naturally ventilated greenhouse in Zimbabwe using a dynamic greenhouse climate model. *Solar Energy, 91*, 381–393. https://doi.org/10.1016/j.solener.2012.09.010.

Nebbali, R., Roy, J. C., & Boulard, T. (2012). Dynamic simulation of the distributed radiative and convective climate within a cropped greenhouse. *Renewable Energy, 43*, 111–129. https://doi.org/10.1016/j.renene.2011.12.003.

Peifer, M., & Timmer, J. (2007). Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET Systems Biology, 1*(2), 78–88. https://doi.org/10.1049/iet-syb.

Pérez Parra, J., Baeza, E., Montero, J. I., & Bailey, B. J. (2004). Natural ventilation of parral greenhouses. *Biosystems Engineering, 87*(3), 355–366. https://doi.org/10.1016/j.biosystemseng.2003.12.004.

Pieters, J. (1999). Modelling solar energy input in greenhouses. *Solar Energy, 67*(1–3), 119–130. https://doi.org/10.1016/S0038-092X(00)00054-2.

Pieters, J., & Deltour, J. M. (1997a). Influence of condenation and evaporation on the climate regulation of greenhouses. In *Proceedings of clima 2000. Brussels (Belgium)* (pp. 1–20).

Pieters, J. G., & Deltour, J. M. (1997b). Performances of greenhouses with the presence of condensation on cladding materials. *Journal of Agricultural and Engineering Research, 68*(2), 125–137. https://doi.org/10.1006/jaer.1997.0187.

Roy, J. C., Boulard, T., Kittas, C., & Wang, S. (2002). Convective and ventilation transfers in greenhouses, Part 1: The greenhouse considered as a perfectly stirred tank. *Biosystems Engineering, 83*(1), 1–20.

del Sagrado, J., Sánchez, J. A., Rodríguez, F., & Berenguel, M. (2016). Bayesian networks for greenhouse temperature control. In *Journal of applied logic* (Vol. 17, pp. 25–35). Elsevier B.V. https://doi.org/10.1016/j.jal.2015.09.006.

Stanghellini, C., van't Ooster, B., & Heuvelink, E. (2019). *Greenhouse horticulture - technology for optimal crop production*. Wageningen Academic Publishers.

Strebel, O. (2013). A preprocessing method for parameter estimation in ordinary differential equations. *Chaos, Solitons and Fractals, 57*, 93–104. https://doi.org/10.1016/j.chaos.2013.08.015.

Taki, M., Ajabshirchi, Y., Ranjbar, S. F., Rohani, A., & Matloobi, M. (2016). Heat transfer and MLP neural network models to predict inside environment variables and energy lost in a semi-solar greenhouse. *Energy and Buildings, 110*, 314–329. https://doi.org/10.1016/j.enbuild.2015.11.010.

Torreggiani, D., Bonora, F., Tassinari, P., Benni, S., & Barbaresi, A. (2016). Efficacy of greenhouse natural ventilation: Environmental monitoring and CFD simulations of a study case. In *Energy and buildings* (Vol. 125, pp. 276–286). Elsevier B.V. https://doi.org/10.1016/j.enbuild.2016.05.014.

Van Henten, E. J. (1994). *Greenhouse climate management: An optimal control approach*.

Yang, J., Chen, J., Shen, Z., Xu, F., Zhang, L., & Zhao, J. (2015). Energy demand forecasting of the greenhouses using nonlinear models based on model optimized prediction method. In *Neurocomputing* (Vol. 174, pp. 1087–1100). Elsevier. https://doi.org/10.1016/j.neucom.2015.09.105.

Yang, C.-K., Chu, C.-R., Lan, T.-W., Tasi, R.-K., & Wu, T.-R. (2017). Wind-driven natural ventilation of greenhouses with vegetation. In *Biosystems engineering* (Vol. 164, pp. 221–234). Elsevier Ltd. https://doi.org/10.1016/j.biosystemseng.2017.10.008.