Available online at www.sciencedirect.com

ScienceDirect

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

Research Paper

Fruit development modelling and performance analysis of automatic greenhouse control

Wouter J.P. Kuijpers [a,*], Duarte J. Antunes [a], Silke Hemming [b], Eldert J. van Henten [c], Marinus J.G. van de Molengraft [a]

[a] Control Systems Technology Group, Eindhoven University of Technology, Eindhoven, 5600 MB, the Netherlands
[b] Greenhouse Horticulture, Wageningen University & Research, Wageningen, 6708 PB, the Netherlands
[c] Farm Technology Group, Wageningen University & Research, Wageningen, 6708 PB, the Netherlands

ARTICLE INFO

Article history:
Received 22 January 2021
Received in revised form
1 May 2021
Accepted 1 June 2021
Published online 30 June 2021

Keywords:
Fruit development
Optimal control
Greenhouse climate control

This paper presents a receding horizon optimal control (RHOC) method with an economic objective function for balancing the cost of resources (resource use × cost) with income through yield (yield × product price). This paper considers the two elements that determine the income through yield. The first element is yield and associated fruit development. A new, computationally viable, approach to model the income through yield is proposed and its prediction accuracy with respect to the original model is evaluated. The new approach employs a model that predicts at each time step, the future income through yield based on the assimilates partitioned to the fruits at the current time step. Simulations suggest that the assumptions made to arrive at the model for the new approach, do not significantly affect the accuracy of the predictions. The second element considered in this paper is the product price and the uncertainty inherent in its forecasts. Historical product price data are used to generate artificial product price forecasts. An uncertainty analysis, in combination with the artificial product price forecasts, showed that the product price forecast error does not considerably affect the optimised control strategy. Season-wide simulations with RHOC suggest that the product price forecast error does not considerably affect the value of the economic objective function.

1. Introduction

The Netherlands is one of the biggest exporters of vegetables in the world, which, however, comes at a price. In 2018, the Dutch horticultural industry consumed 100.5 *PJ* of energy, of which only 7.4 *PJ* was produced in a sustainable manner. This use of energy resulted in a CO_2 emission of 5.7 *Mt* (Velden & Smit, 2019). The Dutch horticultural industry signed an agreement with the Dutch government to decrease the CO_2 emission and its environmental footprint. Among other innovations, automatic control of the greenhouse is likely to contribute to achieving the goals set in the Dutch agreement and may lead to a more sustainable cultivation worldwide (van Straten & van Henten, 2010).

Various applications of automatic control to the greenhouse system have been presented in the literature. In Kuijpers et al.

* *Corresponding author.*
E-mail addresses: w.j.p.kuijpers@tue.nl (W.J.P. Kuijpers), d.antunes@tue.nl (D.J. Antunes), silke.hemming@wur.nl (S. Hemming), eldert.vanhenten@wur.nl (E.J. van Henten), m.j.g.v.d.molengraft@tue.nl (M.J.G. van de Molengraft).
https://doi.org/10.1016/j.biosystemseng.2021.06.002

Nomenclature

Sub- and Superscripts

$\bullet_d$	discretised variable or function
$\bullet^*$	optimal signal or value

Greek Symbols

α	coefficients in function μ_{fft}
β	coefficients in function μ_{fft}
Γ	non-linear function in fruit development model by Vanthoor
δ	Dirac delta function
Δ	difference
θ_l	lower bound to inequality equations $h(\bullet)$
θ_u	upper bound to inequality equations $h(\bullet)$
Θ	non-linear function in fruit development model by Vanthoor
$\mu_{fft}(\bullet)$	average yearly product price trend, €.kg^{-1}
Ξ	impulse response function of a fruit development model, s^{-1}
Σ_x	reference to model component x
τ	variable of integration
τ_l	sampling interval of signals, s
Φ	phase of sinusoids in function μ_{fft}
ω	frequency of sinusoids in function μ_{fft}

Alphabetical Symbols

c	vector of input unit prices
$\widehat{c}$	vector of input unit prices, with product price forecast
c_{frt}	product price, €.kg^{-1}
$\widehat{c}_{frt}$	product price forecast, €.kg^{-1}
$\bar{c}_{frt}$	current product price, €.kg^{-1}
d	uncontrollable inputs to the greenhouse system
$\mathbb{D}$	number of development stages
$FGP(\bullet)$	fruit growth period function, s
$h(\bullet)$	inequality constraint functions
i	integer variable
j	integer variable
k	integer variable
$l(\bullet)$	operational return of the greenhouse system, €.m^{-2}
$l_{frt}(\bullet)$	income through yield, €.m^{-2}
$l_{res}(\bullet)$	cost of resources, €.m^{-2}
n_d	number of uncontrollable inputs
n_e	number of inequality constraints
n_u	number of inputs
n_x	number of states
N	prediction horizon
N_{fruit}	number of fruits in development stage, –
$s(\bullet)$	gas use function, $m^3.m^{-2}$
s_k^l	slope parameter in S, (–)
s_s	switching value parameter in S, (–)
s_v	value that determines value of S, (–)
$S(\bullet)$	smoothed if-else function by Vanthoor (2011)
t	continuous time scale
t_{tsm}	fixed time delay in time-shift model, s
t_i^F	delay of development stage i
T_{can}^{24}	24-hr average greenhouse air temperature, °C
u	controllable inputs to the greenhouse system
u_{CO2}	greenhouse CO_2 injection, $g.m^{-2}.s^{-1}$
u_g	controlled inputs to the greenhouse climate model
$\mathbb{U}$	set of admissible values for the inputs
x	state vector of the greenhouse system
x_t	initial state
x_{frt}	fruit assimilate buffer, C_{fruit} in Vanthoor (2011), $kg.m^{-2}.s^{-1}$
$\bar{x}_{frt}$	steady state value of x_{frt}
$\mathbb{X}$	set of admissible values for the states
y_c	effect of the crop on the greenhouse climate
y_g	effect of the greenhouse climate on the crop
y_B	the assimilates partitioned to the fruits, MC_{BufFrt} in Vanthoor (2011), $kg.m^{-2}.s^{-1}$
y_F	fruit harvest, MC_{FrtHar} in Vanthoor (2011)
$\check{y}_F(t\|t_1)$	isolated contribution of assimilates partitioned at t_1, $kg.m^{-2}.s^{-1}$
y_I	partitioned assimilates that are eventually harvested, $kg.m^{-2}.s^{-1}$
y_M	assimilates used for maintenance respiration, $\sum_{i=1}^{\mathbb{D}} MC_{FrtAir\{i\}}$ in Vanthoor (2011), $kg.m^{-2}.s^{-1}$

Acronyms

CO_2	carbon dioxide
HPS	high pressure Sodium
FGP	fruit growth period
mVt	modified Vanthoor model
RMSE	root-mean-squared-error
RHOC	receding horizon optimal control
SRHOC	receding horizon optimal control with a prediction horizon considerably shorter than the FGP
tsm	time-shift model
Vt	Vanthoor model

(2021), Ramírez-Arias et al. (2012), Seginer et al. (2018), Tap (2000), van Beveren et al. (2015) and Xu et al. (2018) optimal control was used to control the greenhouse system. In RHOC, a subset of optimal control, the control strategy computed by the controller is the result of an objective function optimisation over a future time interval. The objective function can represent various aims, e.g. maximising economic return or yield. In greenhouse climate control based on RHOC with an economic objective function, the cost of resources (resource use × cost) is balanced with the income through yield (yield × product price). The research described in this paper focuses on two elements of the economic objective function. The first element considered is the modelling of yield and associated fruit development. The second element considered is the product price and the uncertainty inherent in its forecasts. Both contributions are introduced in the next two sections.

1.1. *Fruit development modelling*

To be able to optimise the control trajectories for the greenhouse with respect to economic return, a model is required to predict the effect of the control inputs on the income through

yield. The control inputs typically affect yield through the constituted greenhouse climate. Yield is the result of fruit development. Here, we compare various approaches to model fruit development for RHOC.

The fast timescales of the greenhouse climate dynamics (order of tens of minutes) are considerably different from the timescales of the crop development dynamics (order of weeks) (van Straten & van Henten, 2010). The inputs to a system characterised by a slow timescale affect the output of the system over an extended time interval. Thus, to be able to balance resource use, which mainly results from processes with fast timescale, and income through yield, which mainly results from processes with slow timescales, a long prediction horizon is required. An extended prediction horizon leads to more decision variables and thus to an increased computational complexity of the optimisation problem. Consequently, forecasts for product price and e.g. prevailing weather will have to be extended, increasing the uncertainty.

A control approach that considers both timescales is referred to as an integrated approach in van Straten and van Henten (2010). In the literature, various integrated control approaches for the greenhouse system have been presented. In Seginer and Ioslovich (1998), a long prediction horizon was used in combination with a model with a single state. A long prediction horizon in combination with more elaborate models is seldomly used due to its computational complexity. In van Henten (1994), a two-timescale approach was presented to manage the computational complexity of the optimisation problem in a lettuce greenhouse system. The first step of this two-timescale approach aims to optimise the crop state trajectories throughout the growing season, focusing on the slow timescales. The second step aims to optimise the control inputs using the computed seasonal trajectories for crop state and costate. In Tap (2000), the approach was applied to a tomato greenhouse both in simulation as well as in a physical experiment in a greenhouse. In Xu et al. (2018), the two-timescale approach was applied to a simulated lettuce greenhouse. The two-timescale approach is a computationally feasible integrated control approach. The latter approach requires an expression for the (quasi-)steady state of the greenhouse climate system. This requirement is a potential drawback as an analytical expression for the (quasi-)steady state cannot always be derived for non-linear systems (Pavlov et al., 2013). Other expressions for the (quasi-)steady state such as tabular data may lead to discontinuities in the optimisation problem (Betts, 1998). Most hierarchical approaches to greenhouse climate control (Ramírez-Arias et al., 2012) employ a simplified greenhouse climate model to compute climate setpoints. It is, however, unknown to what extent the use of these simplified models affects the performance of the controlled greenhouse system as compared to the use of the original model.

Fruit development is one of the mechanisms which is characterised by a slow timescale. The time from fruit set to fruit harvest is referred to as the fruit growth period (FGP). The FGP is in the range of 40 *days* to 60 *days* (Koning, 1994). Upon partitioning of an assimilate to a fruit, it might take up to the FGP before the respective fruit has ripened enough for harvest. Therefore, instead of calculating the income through yield based on fruit harvest, Kuijpers et al. (2021) and Seginer et al. (2018) calculated the income through yield based on the assimilates partitioned to the fruits. To account for the inherent delay in fruit development, Kuijpers et al. (2021) sold the assimilates partitioned at the current time instance for the product price approximately half FGP later. The latter approaches use simplified models of fruit development and, thus, suffer from a decreased accuracy in predicting income through yield compared to more elaborate models as presented in (Koning, 1994; Vanthoor, 2011). On the other hand, the simplified models do allow for the application of RHOC with a short prediction horizon (e.g. 3 *days*) as presented in Kuijpers et al. (2021). The integration of these models results in a computationally simple optimisation problem with a short prediction horizon. Here, RHOC with a prediction horizon considerably smaller than the FGP is referred to as SRHOC.

The contribution of this research to the modelling of fruit development is two-fold:

- A new fruit development model is proposed, which is based on a generalisation of the work presented in Kuijpers et al. (2021) and Seginer et al. (2018). The new model provides a more accurate description of fruit development than that in Kuijpers et al. (2021) and Seginer et al. (2018) and enables the use of SRHOC as it does not have slow time constants.
- The new model is compared in simulation to that by Kuijpers et al. (2021) to evaluate the effect of the improved accuracy on the performance of the greenhouse system. With this comparison, the suggested performance increase of RHOC, as stated in Kuijpers et al. (2021), can be refined to a more realistic value.

1.2. *Produce price forecasts*

The approach underlying the new fruit development model, as well as any other aforementioned approach, requires a forecast of the product price at the time the fruit is harvested. The product price, nowadays, is determined through price agreements between growers and purchasers mediated through the auction house (Bunte et al., 2009, pp. 1–72). The product price is based on the evolution of market supply and demand. The evolution of market supply and demand are partially based on a seasonal effect, i.e. in summer more tomatoes are produced and supplied to the market. Because of the many (unpredictable) effects influencing market supply and demand, accurate forecasts of the product price do not exist. van Henten (1994) analysed four years of historical market price data for lettuce. The analysis showed that the product price could be predicted using temporal correlations in the data. Most growers, however, do not use product price forecasting services but rely on relevant experience from previous seasons. A product price forecast will seldomly match the actual realisation of the product price, introducing a product price forecast error whose value is uncertain. The effect of the product price forecast error on the performance of a controlled greenhouse system has not been addressed in literature.

This research contributes to the understanding of the effect of product price forecast errors on the performance of the controlled greenhouse system through:

- An analysis of historical product price data to find average seasonal trends in the product price.
- An evaluation of the effect of the product price forecast error on the performance of the greenhouse system. This evaluation uses artificial product price forecasts based on the average seasonal trends in the product price, and the new fruit development model.

1.3. *Paper structure*

The remainder of the paper is organised as follows. Section 2 presents the greenhouse control problem, the proposed fruit development model and the artificial price forecasts. The various simulation studies and corresponding results are presented in Section 3. The results of the simulation studies are discussed in context of the above contributions in Section 4. Concluding remarks and indications for future research are presented in Section 5.

2. Materials & methods

The research presented here builds upon the greenhouse control problem as presented in Kuijpers et al. (2021). Relevant parts of this control problem are presented in Subsection 2.1. In Subsection 2.2, the approach proposed in Kuijpers et al. (2021) and Seginer et al. (2018) is generalised to arrive at a fundamentally different way to model fruit development which enables the use of SRHOC. Subsection 2.3 details modifications to the model by Vanthoor (2011) according to the specifications set out in Subsection 2.2. The analysis of the historical product price data and the synthesis of artificial product price forecasts are presented in Subsection 2.4.

2.1. *Greenhouse control problem*

The controlled greenhouse system is graphically represented by the block diagram in Fig. 1. The model of the greenhouse system is composed of the energy management system ΣE, greenhouse climate system model and lighting system model ΣG and crop growth and transpiration model ΣC. The fruit development model ΣFD, a key model component for the purpose of this paper, is depicted as a separate block. The inputs to the fruit development model block are the greenhouse climate y_g and the assimilates partitioned to the fruit buffer[1] $y_B \in \mathbb{R}$. The greenhouse system is modelled with a state-space representation with states $x_d \in \mathbb{R}^{n_x}$, controllable inputs $u_d \in \mathbb{R}^{n_u}$ and uncontrollable exogenous inputs $d \in \mathbb{R}^{n_d}$. Further details on the model used in this paper can be found in Kuijpers et al. (2021). The simulation studies presented here are based on a greenhouse with high-pressure sodium (HPS) lamps.

In the simulation studies presented in this paper, the control inputs of the greenhouse system u_d resulted from an optimisation problem solved by controller Σ_M. The controller aimed to optimise the operational return $J \in \mathbb{R}$ ($€.m^{-2}$), over a finite receding horizon. This horizon starts at time instance k and subsequent time instances in the horizon are denoted by $j \in \{0, \dots, N\}$, where N is the length of the prediction horizon. The operational return constitutes a purely economic objective function in which the inputs are multiplied with cost/return of a unit input $c(t) \in \mathbb{R}^{n_u}$, i.e. $c_i(t)u_i(t)$ represents the contribution of input $i \in \{1, \dots, n_u\}$ to the operational return and is expressed in monetary units. The input vector u_d includes elements with a negative contribution to the return (costs) such as gas use or electricity use. The input vector u_d also includes elements with a positive contribution (profit) i.e. fruit harvest and selling electricity. The overall cost, referred to as the operational return $l_d(\cdot) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_u} \to \mathbb{R}$, is $l_d(u_d, c) = c^T u_d$. Although the weights c, except for the product price $c_{frt} \in \mathbb{R}$, were chosen constant, they can, in general, vary with time. Here, $\hat{c}_{frt}(j|k) \in \mathbb{R}$ represents a product price forecast published at time instance k with lead time j, forecasting a future time instance. $j + k$

The discrete-time states of the greenhouse system $x_d \in \mathbb{R}^{n_x}$ affect the optimisation of the controllable inputs u_d through the constraints. The controllable inputs to the greenhouse system u_d can be updated every 15 min, $\tau_l = 15 \cdot 60 = 900$ s and are held constant in between samples. In Kuijpers et al. (2021), a discretization period of 225 s was found to be sufficiently short for the discretised system to accurately represent the continuous behaviour of the system. For the sake of computational efficiency, these inputs are discretised using a zero-order hold with sampling time τ_l. The optimisation problem aims to find u_d^* by maximising the operational return within the feasible region outlined by the constraints, summarized in the following optimal control problem

$$u_d^* = \underset{u_d(\cdot|k)}{\operatorname{argmax}} \sum_{j=0}^{N} l_d(u_d(j|k), \hat{c}(j|k)), \tag{1}$$

Subject to:

$$\begin{aligned} &x_d(j+1/k) = F(x_d(j/k), u_d(j/k), d(j/k)), \\ &(x_d(j/k), u_d(j/k)) \in X \times U \\ &\theta_l \le h(x_d(j/k), u_d(j/k)) \le \theta_u, \quad \forall j \in \{0, \dots, N\} \\ &x_d(0/k) = x_t \end{aligned}$$

The state of the system at the present time is represented by $x_t \in \mathbb{R}^{n_x}$. Vector $x_d(j|k)$ is the predicted[2] state at future time instance $j + k$ computed at time instance k. The inputs to the greenhouse at time instance $j + k$ predicted at time instance k are represented by $u_d(j|k)$. The forecast of the uncontrollable inputs to the greenhouse $d(j|k) \in \mathbb{R}^{n_d}$ encompasses the relevant weather variables i.e. global radiation, outside air temperature, outside air CO_2 concentration, outside air humidity and wind speed. The weather forecast $d(j|k)$ input to (1) is equal to the realised weather. The dynamical model $F : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_d} \to \mathbb{R}^{n_x}$ represents Σ_E, Σ_G, Σ_C and Σ_{FD} and provides a mapping from inputs and states to the states τ_l into the future. Further details on model F can be found in Kuijpers et al. (2021). It is implicit in (1) that the controller is assumed to have full state information.

The optimisation problem in (1) was solved using the nonlinear optimisation software package IPOPT (Wächter & Biegler, 2006), with linear solver MA57 from HSL (HSL, 2019).

[1] The notation and naming of the signals in the model by Vanthoor (2011), can be found in the respective entry in the Nomenclature.

[2] the notation ·(k|t) denotes the variable ·(k+t) predicted at time instance t.

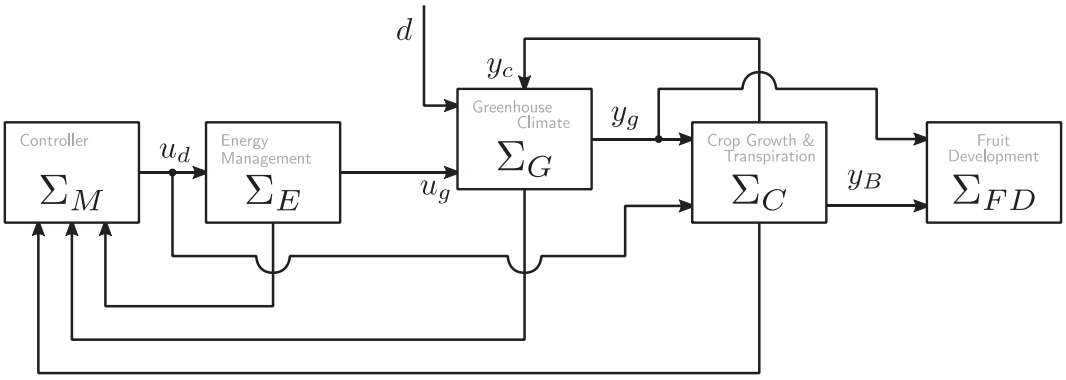

Fig. 1 – Block diagram representation of the greenhouse control system. The system is composed of controller Σ_M, energy management system Σ_E, greenhouse climate and lighting system model Σ_G and crop growth and transpiration model Σ_C and fruit development model Σ_{FD}. The output of the controller is denoted by u_d. The elements of u_d input to Σ_C encompass the harvest of fruits and leaves. The controllable inputs to the greenhouse climate model, e.g. ventilation, screen deployment, heating, CO_2 injection are denoted by u_g. The uncontrollable inputs to the greenhouse climate model are denoted by d, the outside weather. Variables y_g and y_c denote the effect of the greenhouse climate on the crop (temperature, CO_2 concentration, radiation and relative humidity) and the effect of the crop on the greenhouse (assimilation and transpiration), respectively. The assimilates partitioned to the fruit buffer are represented by y_B.

The IPOPT solver was interfaced using CasADI (Andersson et al., 2019). To improve the convergence of the solver, CasADI provides it with Hessian and Lagrangian information of the optimisation problem in (1). The latter requires the functions F and l_d in (1) to be twice continuously differentiable in order for the Hessian to be defined for any given input.

The research presented in this paper focuses on the model for the income through yield $l_{frt}(\cdot) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_u} \to \mathbb{R}$. This subsection continues with a more detailed description of the income through yield. For multiple-harvest crops such as tomatoes, income through yield is obtained multiple times throughout the prediction horizon. For tomato, the income through yield is therefore often included in the running cost function (Ramírez-Arias et al., 2012; Seginer et al., 2018). The operational return can thus be decomposed according to

$$l_d(u_d, c) = l_{frt}(u_d, c) - l_{res}(u_d, c), \tag{2}$$

where $l_{res}(\cdot) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_u} \to \mathbb{R}$ represents the cost of resources,[3] i.e. the amount of resources used times the cost per resource. Fruits, like tomatoes, require a certain amount of time to ripen. The time from fruit set to fruit harvest is referred to as the FGP. The FGP is affected by temperature and by light through photosynthesis. The cost of resources $l_{res}(u_d, c)$ is balanced once the fruit is sold, which can, thus, take up to 60 *days*. For example, by turning on the artificial lighting, the resource costs increase as a result of increased gas use or electricity use. These costs will only be fully balanced when all the fruits that consist of assimilates that were affected by this change in control strategy have been harvested and sold. Given this FGP, one approach for implementing this optimisation problem would be to use a prediction horizon of 60 *days*. Given the continuous nature of the investments in assimilates using resources, this, would already be an approximation of the solution of the control problem for the whole production period. Yet even this horizon of 60 *days* is not feasible for most numerical solvers.

[3] selling electricity to the grid is included as negatively contributing to l_{res}.

2.2. *Fruit development modelling*

In this subsection, a fundamentally different approach to fruit development modelling is presented. This subsection starts with an introduction of the new approach and a comparison with respect to the approach used in most literature. As opposed to the discrete timescale used in (1), we treat this approach in the continuous time as most models are. The employed discretization method is detailed in Kuijpers et al. (2021).

Let t denote a continuous timescale. On timescale t, let t_1 denote the specific time instant at which the optimisation algorithm is solved. In literature, the income through yield is typically described based on a prediction of the fruit harvest $y_F \in \mathbb{R}$. This prediction describes the amount of assimilates that will be harvested at a time instant t_2. Let y_B denote the assimilates partitioned to the fruit buffer, see Fig. 1. The part of the assimilates that are partitioned at the current time instant and that is eventually harvested is $y_I = y_B - y_M \in \mathbb{R}$ in which $y_M \in \mathbb{R}$ denotes the assimilates used for maintenance respiration, these do not contribute to the eventual fruit weight. As the output of the fruit development model is characterised by a slow timescale, the fruit harvest at t_2, i.e. $y_F(t_2)$, is the result of y_I over a relatively long period of time. This becomes apparent from the convolution

$$y_F(t_2) = \int_0^{FGP} \Xi(t_2 - \tau) y_I(\tau)\, d\tau, \tag{3}$$

As $y_F(t_2)$ is a function of input $y_I(t_i)\ \forall t_i \in [t_2 - FGP, t_2]$ Note that $\Xi(t_2 - \tau) = 0\ \forall\ \tau > FGP$, hence the upper limit of the convolution integral was changed to FGP. The proposed approach requires the existence of a unique impulse response Ξ, the implications of this requirement are discussed later. As $y_F(t_2)$ is the result of input y_B over the interval $[t_2 - FGP, t_2]$, the optimisation problem in (1) requires a prediction horizon N the length of which exceeds FGP.

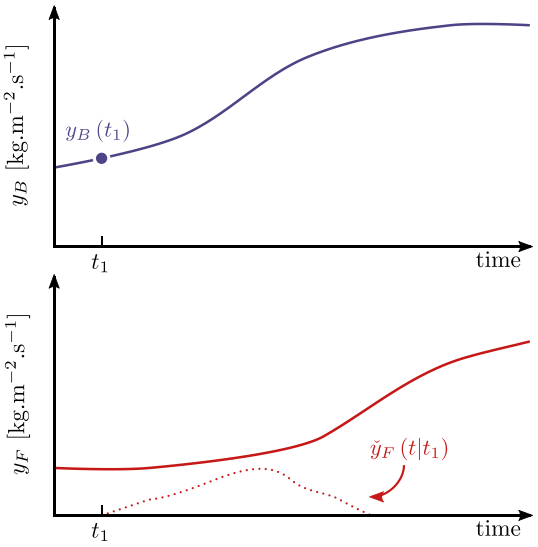

Fig. 2 – Illustration of the assimilates partitioned to the fruit buffer y_B (top), and resulting fruit harvest y_F (bottom) predicted by a fruit development model. The dashed line in the bottom graph indicated by $\check{y}_F(t|t_1)$ indicates the fruits which are harvested due to assimilates partitioned at t_1, i.e. $y_B(t_1)$ as indicated in the top graph.

The underlying principle of the fundamentally different approach to fruit development modelling is as follows. Let $\check{y}_F(t|t_1) \in \mathbb{R}$ denote the isolated contribution of $y_B(t_1)$ to the fruit harvest $y_F(t)$, hence

$$\check{y}_F(t|t_1) = \lim_{||\Delta t|| \to 0} \Xi(t - t_1) y_I(t_1)\ \Delta t. \tag{4}$$

In (4), $\check{y}_F(t|t_1)$ is a function of only $y_I(t_1)$ and thus y_B at t_1. Variable $\check{y}_F(t|t_1)$ represents the fruits that are harvested at t due to partitioned assimilates at t_1. Figure 2 illustrates this step by visualising the part of y_F that is $\check{y}_F(t|t_1)$. The proposed approach models the income through yield as a function of the isolated contribution of the assimilates partitioned at time instance t_1, $\check{y}_F(t|t_1)$, multiplied by the product price forecast as

$$l_{frt}(t_1) = \int_0^{FGP} \check{y}_F(\tau|t_1) \cdot \widehat{c}(\tau|t_1)\ d\tau. \tag{5}$$

Note that, even though the integral in (5) is also over the interval $[0,\ FGP]$, l_{frt} in (5) is only a function of $y_B(t_1)$. As this approach to model the income through yield is only a function of the input y_B at t_1 it does not require a prediction horizon which exceeds FGP.

The income through yield in (5) represents the general form of the approach, i.e. $\Xi(\cdot)$ is used to denote the impulse response function of any linear time-invariant fruit development model. The method in Kuijpers et al. (2021) and Seginer et al. (2018), here referred to as time-shift model (tsm), can be written in the form of (5). The tsm predicts the economic contribution of the assimilates partitioned to the fruit buffer y_B by assuming that these will be harvested after a fixed time $t_{tsm} \in \mathbb{R}^{\geq 0}$. The impulse response function of the tsm represents a pure delay process $\Xi(t) = \delta(t - t_{tsm})$, i.e.

$$l_{frt}(t_k) = \int_0^{FGP} \delta(\tau - t_{tsm}) \cdot y_I(t_k) \cdot \widehat{c}_{frt}(\tau|t_k)\ d\tau, \tag{6}$$

where $\delta(\cdot) : \mathbb{R} \to \mathbb{R}$ denotes the Dirac delta function. Note that the combination of the Dirac delta function and the integral is included to show the similarity between (6) and (5).

The assumption that underlies the tsm and the explanation at the start of this subsection is that the fruit harvest y_F can be described as the sum of the contributions of the individual assimilates partitioned at respective time instances. In other words, the impulse response function $\Xi(\cdot)$ in (3) exists. The principle of superposition holds for linear models, when time-invariant the model has a unique impulse response. The tsm consists of a single delay operator and is therefore a linear time-invariant model. In the next subsection, the model by Vanthoor (2011) is modified to arrive at a linear time-invariant model.

2.3. *Modified vanthoor model*

Before discussing the assumptions and modifications to the model by Vanthoor (2011) to arrive at a linear time-invariant model, first a short introduction of model is presented. Fruit development was modelled by Vanthoor (2011) using a fixed boxcar train structure (Rabbinge & Ward, 1989) based on $\mathbb{D}$ distinct fruit development stages. Each stage i contains an assimilate buffer $x_{frt\{i\}}$ which represents the assimilates of fruits in development stage i and buffer $N_{fruit\{i\}}$ which represents the number of fruits in development stage i. The fruit development model by Vanthoor is represented by a block diagram in the top panel of Fig. 3. Note that, for the sake of clarity, Fig. 3 focuses on the assimilate buffers in the model, the fruit number buffers $N_{fruit\{i\}}$ are not shown. The assimilates partitioned to the fruits y_B are distributed over the development stages through gains Θ_j. The assimilates flow through the subsequent stages until the fruit is harvested after stage $\mathbb{D}$. The gains Γ_i control the flow from buffer to buffer. The gains Θ_i and Γ_i are non-linear functions of $N_{fruit\{i\}}$ and the 24 *hr* average temperature T^{24}_{can}, for the sake of clarity these dependencies are not shown in Fig. 3 and in the following elaboration. The fruit development model by Vanthoor in state space representation is given by

$$\begin{cases} \begin{bmatrix} \dot{x}_{frt\{1\}}(t) \\ \vdots \\ \dot{x}_{frt\{\mathbb{D}-1\}}(t) \\ \dot{x}_{frt\{\mathbb{D}\}}(t) \end{bmatrix} = \begin{bmatrix} -\Gamma_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -\Gamma_{\mathbb{D}-1} & 0 \\ 0 & \cdots & \Gamma_{\mathbb{D}-1} & -\Gamma_{\mathbb{D}} \end{bmatrix} \begin{bmatrix} x_{frt\{1\}}(t) \\ \vdots \\ x_{frt\{\mathbb{D}-1\}}(t) \\ x_{frt\{\mathbb{D}\}}(t) \end{bmatrix} - \begin{bmatrix} y_{M\{1\}} \\ \vdots \\ y_{M\{\mathbb{D}-1\}} \\ y_{M\{\mathbb{D}\}} \end{bmatrix} + \begin{bmatrix} \Theta_1 \\ \vdots \\ \Theta_{\mathbb{D}-1} \\ \Theta_{\mathbb{D}} \end{bmatrix} y_B, \\ l_{frt}(t) = \Gamma_{\mathbb{D}} x_{frt\{\mathbb{D}\}}(t) c_{frt}(t) \end{cases} \tag{7}$$

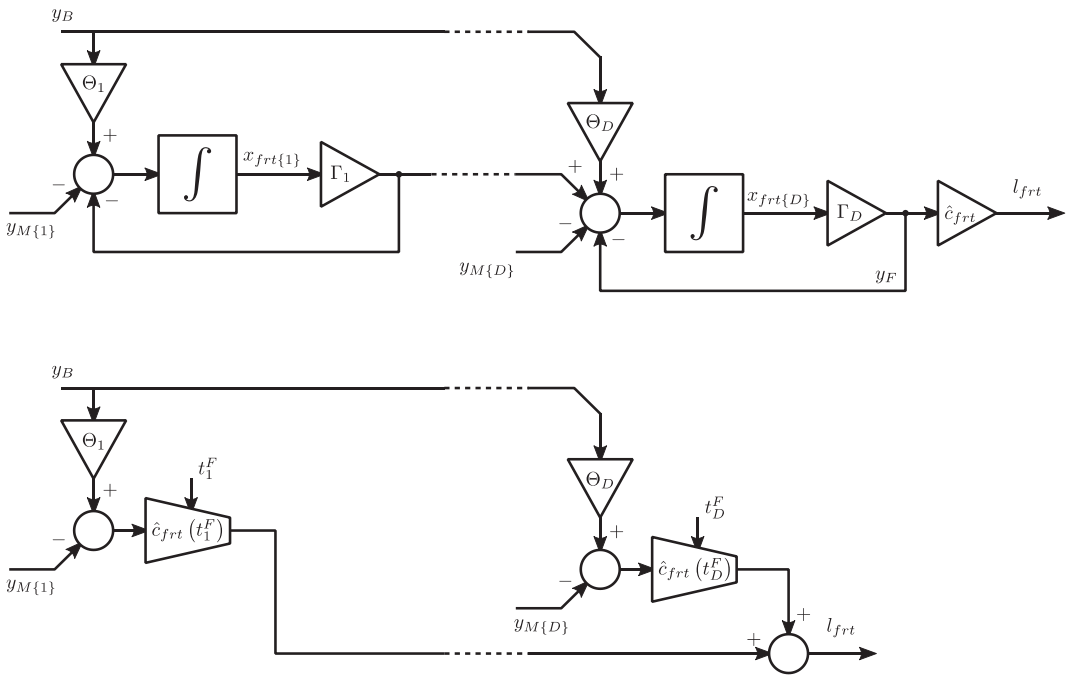

Fig. 3 – (top) A block diagram representation of the fruit development model by Vanthoor (2011). The assimilates that are partitioned to the fruits y_B flow to various development stages, the distribution over the development stages is controlled by factors Θ_i $\forall i \in \{1, \ldots, \mathbb{D}\}$. The assimilates required for fruit maintenance respiration $y_{M\{i\}}$ are withdrawn from each fruit buffer $x_{frt\{i\}}$. The assimilates flow from buffer to buffer controlled by Γ_i. To arrive at the resulting crop price, the harvested assimilates y_F are multiplied by the forecasted product price $\hat{c}_{frt}$. (bottom) a block diagram representation of the modification to the fruit development model by Vanthoor, the block with $\hat{c}_{frt}(t_i^F)$ represents a multiplication of the input with the product price $\hat{c}_{frt}$ a period of length t_i^F into the future.

where $y_{M\{i\}} \in \mathbb{R}^{\mathbb{D}}$ represents the maintenance respiration of stage i. As this model provides a relation between the assimilates partitioned to the fruits and income through yield l_{frt}, it can be substituted for Σ_{FD} in Fig. 1. The Vanthoor model in (7) is, because of non-linear functions Θ_i and Γ_i, non-linear and can therefore not be substituted for Ξ in (4). In the remainder of this subsection, the modifications to the model in (7) are presented, to arrive at a linear time-invariant model.

The fixed boxcar train approach in (Rabbinge & Ward, 1989), applied to fruit development by Vanthoor (2011), is an approach to represent a time delay response of a system. In Vanthoor (2011), this delay was modelled as a function of the development stage the assimilates are partitioned to and the 24 *hr* average temperature T_{can}^{24} according to

$$t_i^F = FGP\left(T_{can}^{24}\right) \cdot \left(1 - \frac{(i-1)+0.5}{\mathbb{D}}\right). \tag{8}$$

Vanthoor models the FGP as a function of T_{can}^{24}, hence $FGP(T_{can}^{24})$. Assimilates partitioned to a higher development stage will be harvested sooner, hence a shorter delay t_i^F is induced. This is supported by the observation that the model in (7) is a chain of integrators and that the eigenvalues of the state matrix in (7) correspond to the delay times t_i^F. Instead of using a fixed boxcar train approach, we propose to directly time-shift the assimilates partitioned to stage i by time delay t_i^F. The fruit harvest can therefore be described as a set of $\mathbb{D}$ parallel delays

$$l_{frt}(t_k) = \sum_{j=1}^{\mathbb{D}} \left(\Theta_j \cdot y_B(t_k) - y_{M\{j\}}(t_k)\right) \cdot \hat{c}_{frt}\left(t_j^F \middle| t_k\right). \tag{9}$$

Note that, as opposed to the description of the tsm in (6), the combination of the Dirac delta function and convolution integral is omitted in (9). The modified Vanthoor model (mVt) in (9) models the fruit harvest as a result of input y_B, at $\mathbb{D}$ distinct time instances t_j^F at which an amount of $\Theta_j \cdot y_B(t_k) - y_{M\{j\}}(t_k)$ for $i \in \{1, \ldots, \mathbb{D}\}$. The impulse response functions of the tsm in (6) and the modified Vanthoor model (mVt) are presented in Fig. 4. From Fig. 4, one can observe how the assimilates are partitioned over the various fruit development stages through gains Θ_j. The mVt is represented as a block diagram at the bottom of Fig. 3. The integrators in the development stages were replaced by blocks where the assimilates, previously input to the integrator, are multiplied by the time-shifted price forecast $\hat{c}_{frt}$. One can observe in the bottom panel of Fig. 3 and from (9) that l_{frt} can be negative for time instances where the amount of assimilates needed for maintenance respiration exceeds that of the assimilates partitioned to the fruit buffers. The negative values of l_{frt} are compensated by positive value at other time instances in the prediction horizon. Two important assumptions are made to arrive at (9). These two assumptions are discussed in the remainder of this subsection. Note that these two assumptions are added to the assumptions on the fruit development process that resulted in the model by Vanthoor.

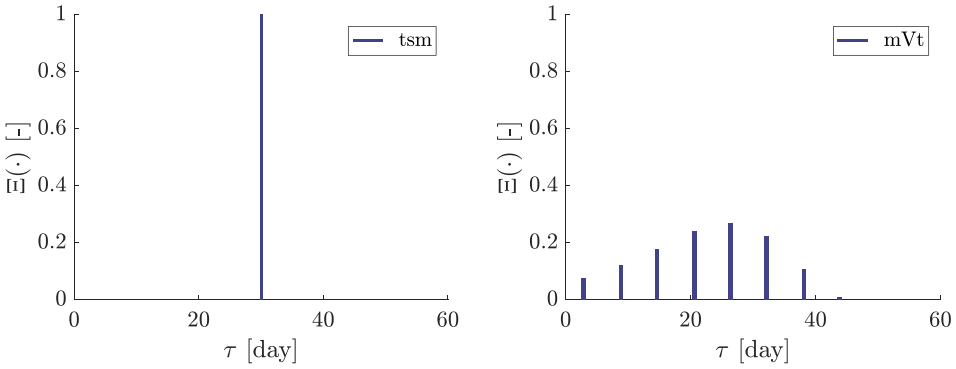

Fig. 4 – Impulse response functions of (left) the time shift model in (6) with $t_{tsm} = 30$ *days* and (right) the modified Vanthoor model in (9) with $T^{24}_{can} = 22\,^{\circ}C$. Both functions are characterised by a combination of impulses.

The first assumption is that the 24 *hr* average temperature T^{24}_{can} remains constant throughout the *FGP*. The model in (9) depends on T^{24}_{can} through t^{F}_{j}. In (9), the value of T^{24}_{can} at t_k is used to calculate l_{frt} and is therefore assumed to remain constant for a period of t^{F}_{j}. The assimilates partitioned at t_k are thus assumed to be unaffected by the temperature after partitioning. If the average temperature realised after partitioning would be 1 °C higher as compared to the value used to calculate t^{F}_{j}, e.g. 21 °C, the resulting error in t^{F}_{j} would be 2.2 *d*. Note that the evolution of $\widehat{c}_{frt}$ is important in evaluating the effect of this assumption, i.e. if $\widehat{c}_{frt}$ is e.g. constant throughout the course of those 2.2 *days*, an error of 1 °C around $T^{24}_{can} = 21\,^{\circ}C$ would yield the same estimate of l_{frt}. At the bottom of Fig. 5, a season long trajectory for T^{24}_{can} is presented for a simulation with HPS lighting in Kuijpers et al. (2021). From Fig. 5, one can observe that in this simulation T^{24}_{can} is constant throughout most of the year.

The second assumption is that the assimilate buffers x_{frt} in (7) and $N_{fruit\{i\}}$ have achieved a steady-state $\overline{x}_{frt} \in \mathbb{R}^{\mathbb{D}}$. As a result of this assumption, the non-linear functions Θ_i and Γ_i are independent of the states of (7). In the top panel of Fig. 5 the resulting state trajectories of x_{fd} are shown for a growing season-wide simulation with the model in (7). The inputs y_B and T^{24}_{can} for this simulation are from Kuijpers et al. (2021). The steady-state value, $\overline{x}_{frt}$ in Fig. 5 results from substitution of the inputs into the expression for the steady-state found by solving $\dot{x}_{frt} = 0$. The generative growth starts at the 11th of January; hence the buffers are empty. One can observe that the buffers achieve a steady state in April, three months after the start of generative growth. The model is therefore considered to provide meaningful predictions roughly 3 *months* after the start of the productive phase. The algorithm in (1) can therefore only be used after this period.

2.4. *Product price and artificial price forecasts*

In the Netherlands, tomatoes are sold through the daily market and through long-term contracts (Bunte et al., 2009, pp. 1–72). The research presented in this paper is valid for tomatoes sold through the daily market. The product price on the daily market is determined through price agreements between growers and purchasers mediated through the auction house (Bunte et al., 2009, pp. 1–72). This price is based on the evolution of market supply and demand. Because of the many (unpredictable) effects influencing market supply and demand, accurate long-term forecasts of the product price do not exist.

As product price forecasts were not available to use, artificial product price forecasts were created to be able to analyse the effect of the product price forecast errors. These forecasts are based on what we consider to be relevant experience of the grower. The relevant experience of the grower consists of the average evolution of the product price throughout the growing season and the average deviation thereof. This subsection continues with an analysis of the historical product prices to obtain these two trends, i.e. the average yearly product price evolution and corresponding deviations thereof.

Figure 6 presents the historical product price data c_{frt} during five years plotted with respect to the day of the year (GFactueel, 2020). This data for the Dutch market consists of five years of daily data for the truss tomato product price at weekdays. The product price for truss tomato is used to match the product price in Kuijpers et al. (2021). One can observe a yearly oscillation which results from the dependency of the product price on supply and demand. In summer the supply is high, hence the product price drops accordingly. A fast Fourier transform was used to identify this yearly oscillation, see Fig. 7. In the frequency spectrum in Fig. 7, one can observe an additional half-yearly oscillation and a quarter-yearly oscillation. The latter are hypothesised to represent the seasonal variations in the product price. The contribution of these three frequencies was used to represent the average yearly trend $\mu_{fft}(t) : \mathbb{R} \to \mathbb{R}$, defined as

$$\mu_{fft}(t) = \alpha + \sum_{i=1}^{3} \beta_i \sin(\omega_i t + \varphi_i), \tag{10}$$

where the values for α, β_i and φ_i result from the fast Fourier transform, results for the magnitude are presented in Fig. 7. The frequency of the first periodic was $\omega_1 = 2\pi \cdot 365^{-1}$, $\omega_2 = 2\omega_1$ and $\omega_3 = 4\omega_1$. The resulting average yearly trend μ_{fft} is presented in Fig. 6. A normally distributed product price was assumed at every day, the standard deviation was determined

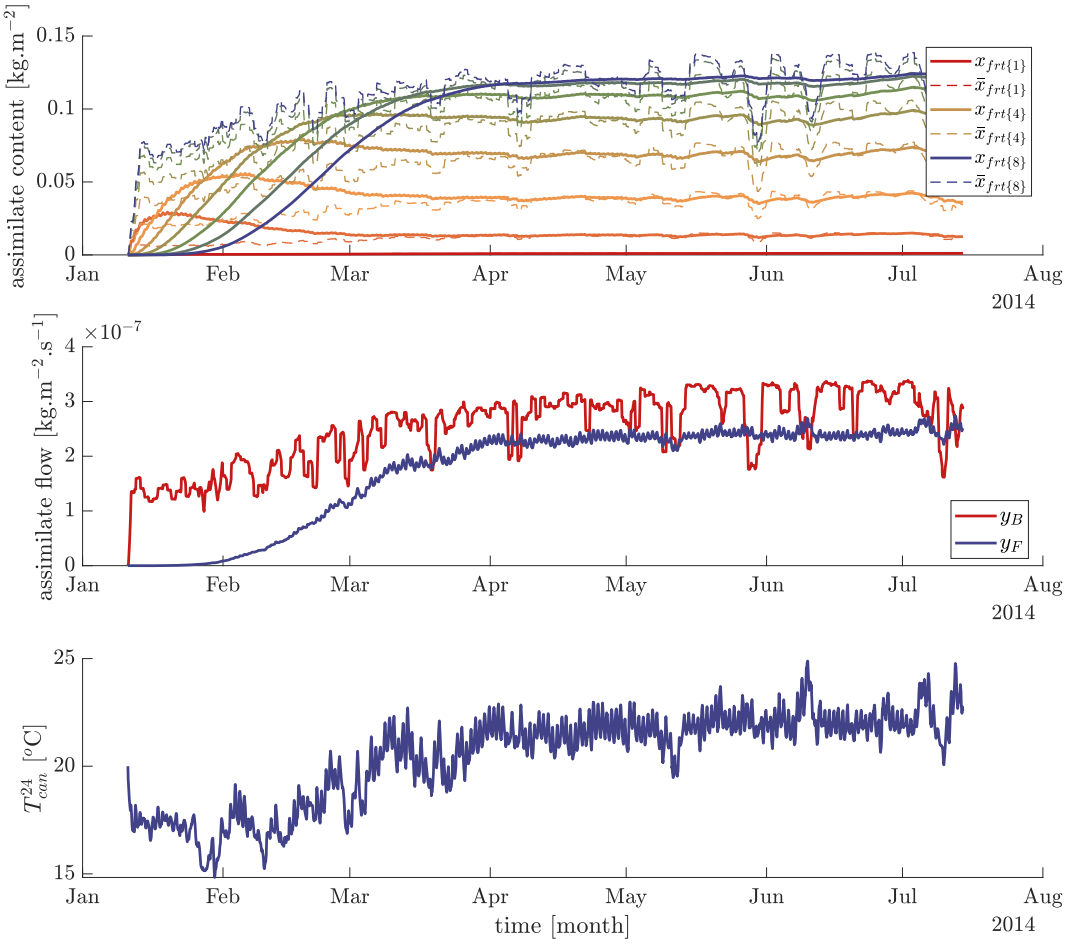

Fig. 5 – Inputs and output of the fruit development model by Vanthoor (2011) with $\mathbb{D} = 8$ using y_B and T^{24}_{can} from a season-wide simulation of an optimally controlled greenhouse with HPS lighting in Kuijpers et al. (2021) (top) assimilate content of $x_{frt\{i\}}$ and the calculated steady-state $\bar{x}_{frt\{i\}}$ (middle) the assimilates partitioned to the fruit buffers y_B and fruit harvest y_F (bottom) the realised T^{24}_{can}. One can observe that the assimilate buffers $x_{frt\{i\}}$ achieve the corresponding steady state $\bar{x}_{frt\{i\}}$ around April. The realised T^{24}_{can} is relatively close to constant throughout the period in which the states are at a steady state value.

by equating the largest difference between c_{frt} and μ_{fft} to 3σ. The fast Fourier transform of the latter signal also revealed three prominent frequencies, a half-yearly, quarter-yearly, and bi-monthly oscillation. In order to model the standard deviation σ_{fft}, a similar structure as (10) is used. The resulting 3σ-bounds are represented by the dashed lines in Fig. 6.

In this research, two different types of artificial forecasts were used to analyse the effect of the product price forecast error. The first type of artificial forecasts represents the experience of the grower, a realistic product price forecast. This first type is based on the belief of the grower that the price will evolve from the current product price $\bar{c}_{frt} \in \mathbb{R}$ to μ_{fft}. The forecasted evolution of the product price was modelled using a smoothed if-else function $S : \mathbb{R}^3 \to \mathbb{R}$, proposed by Vanthoor (2011).

$$S(s_v, s_k^l, s_s) = \frac{1}{1 + e^{s_k^l (s_v - s_s)}}, \tag{11}$$

where $s_k^l(-)$ and $s_s(-)$ represent the slope and the switching value of the smoothed if-else. The lead time of the forecast was substituted for $s_v(-)$ in (11). The smoothed if-else function was chosen because of the fact that its parameters have a clear interpretation and that it provides a smooth description. The latter is desirable for the optimisation algorithm. The first set of artificial forecasts is described by

$$\hat{c}_{frt}(j|t) = \mu_{fft}(t) + \bar{c}_{frt} \cdot S(j, s_k^l, s_s), \tag{12}$$

where j represents the lead time of the price forecast. As $j \to 0$ in (12), $S(j, s_k^l, s_s) \to 1$ as $s_s \gg 0$ and $s_k^l > 0$ resulting in $\hat{c}_{frt}(j|t) = \mu_{fft}(t) + \bar{c}_{frt}$. For $j \to \infty$, $S(j, s_k^l, s_s) \to 0$ and $\hat{c}_{frt}(j|t) = \mu_{fft}(t)$. The value for the slope s_k^l and switching value s_s were randomly selected upon generation of the forecast. A set of price forecasts $\hat{c}_{frt}$ which was generated using (12) and random $s_k^l > 0$ and $s_s \gg 0$, is presented in the right panel of Fig. 8.

A second set of artificial forecasts was designed to evaluate the effect of the product price forecast uncertainty on the optimised control strategy. This product price forecast error was assumed to be zero at the current time instance but assumed to converge to $3\sigma_{fft}$ for increasing lead time. The

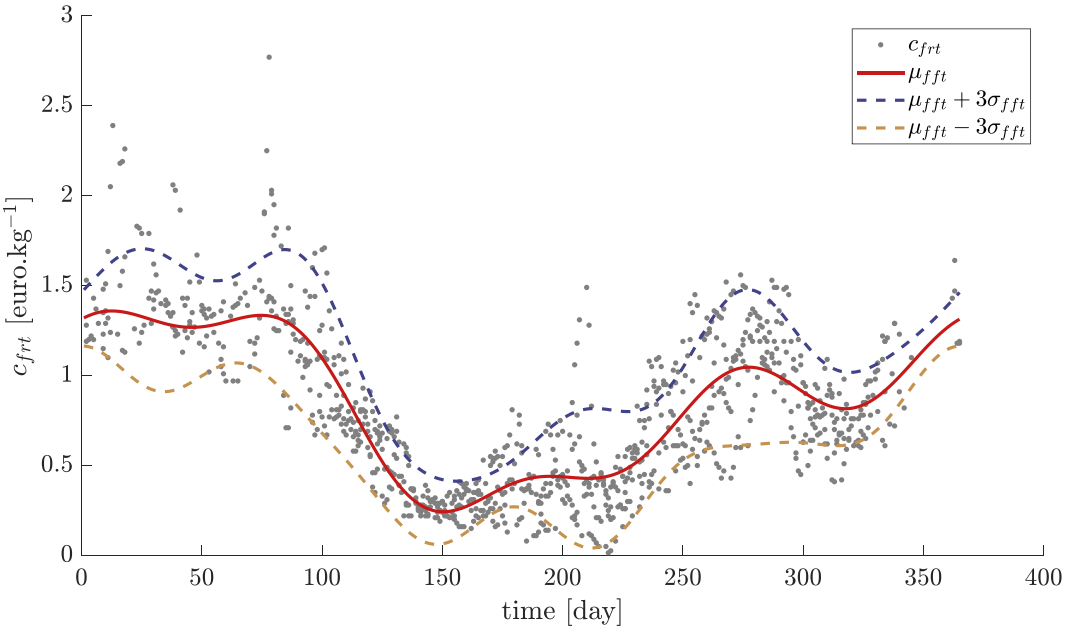

Fig. 6 – Historical product price data c_{frt} of five years plotted with respect to the day of the year. Also, the average yearly trend μ_{fft} and the lower and upper 3σ-bounds are plotted.

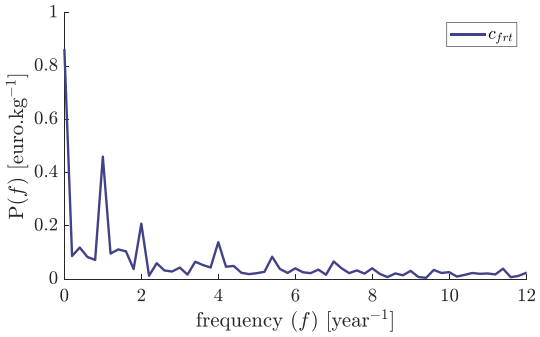

Fig. 7 – Fast Fourier transform of the historical product price data. The frequency is expressed in *year*$^{-1}$, the peak at 1, thus, represents the yearly oscillation in the product price.

latter represents the increasing uncertainty for product price forecasts. The artificial forecasts in the second set are equal to $\mu_{fft}(k)$ for $l = 0$ and converge to the positive or negative 3σ bound for increasing j. This set of artificial forecasts, used in the simulation study presented in Subsection 3.2, is described by

$$\hat{c}_{frt}(j|t) = \mu_{fft}(t) \pm 3\sigma_{fft}(t) \cdot S(j, s_k^l, s_s). \tag{13}$$

As $j \to 0$, $S(j, s_k^l, s_s) \to 0$ as $s_s \gg 0$ and $s_k^l < 0$ resulting in $\hat{c}_{frt}(l|t) = \mu_{fft}(t)$. For $j \to \infty$, $S(j, s_k^l, s_s) \to \pm 1$ and $\hat{c}_{frt}(l|t) = \mu_{fft}(t) \pm 3\sigma_{fft}(k)$. Values for the slope $s_k^l < 0$ and switching value s_s as well as whether the contribution of $3\sigma_{fft}(t)$ is added or subtracted from $\mu_{fft}(t)$ were randomly selected. A set of price forecasts which was generated using (13) and random s_k^l and s_s is presented in the left panel of Fig. 8.

2.5. *Weather data*

In the simulation studies presented here, the realisation of the weather applied to the system d originated from an experiment described in Kempkes et al. (2014), where various energy-saving options in greenhouses were investigated in a Venlow Energy kas located in Bleiswijk, The Netherlands. The data, which consist of outside air temperature T_{out} (°C), outside absolute air humidity H_{out} (g.m^{-3}), outside air CO_2 concentration $CO_{2,out}$(g.m^{-3}), wind speed v_{wind} (m.s^{-1}) and global radiation Q_{sun} (W.m^{-2}), are measured at 5 *min* interval, during the years 2011 to 2014. In the simulation studies here, only year 2014 was used.

3. Results

In the first simulation study in Subsection 3.1, the three fruit development models presented in Section 2, i.e. the tsm, the model by Vanthoor (2011) and the mVt are compared in an open-loop simulation. Subsection 3.2 presents a simulation study to assess the effect of the product price forecast error uncertainty on the optimised control trajectory resulting from (1). Subsection 3.3, presents a season-wide simulation in which the tsm and mVt fruit development models are used in closed-loop, i.e. with the models integrated in (1).

3.1. *Fruit development model comparison*

In this simulation study, the three fruit development models presented in Section 2: the tsm in (6), the model by Vanthoor (2011) in (7) denoted by Vt, and the mVt in (9), were provided with the same inputs in order to compare the resulting predictions.

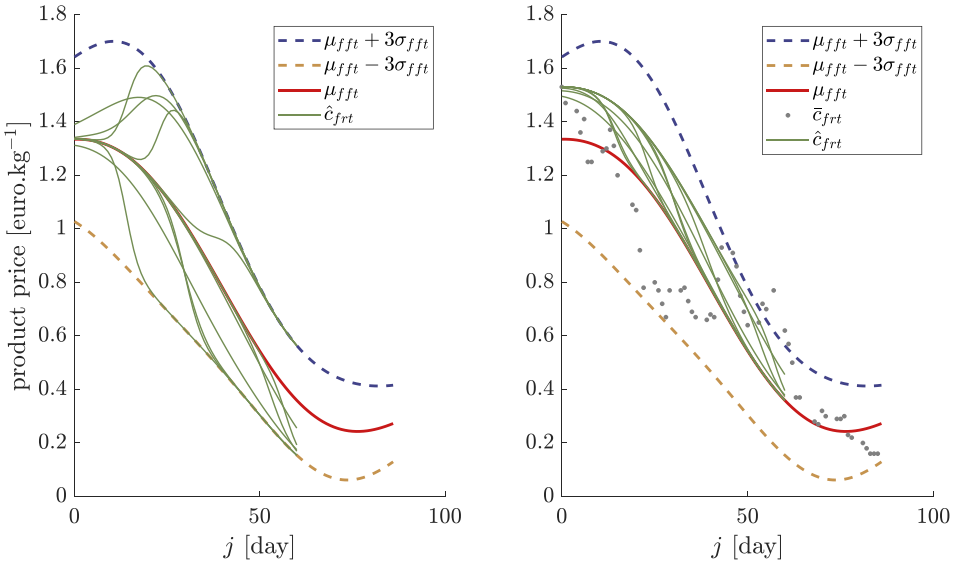

Fig. 8 – The averaged product price μ_{fft} and corresponding 3σ bounds throughout the interval from March 15th to June 9th, (left) the forecasts $\hat{c}_{frt}$, used in the sensitivity analysis, resulting from (13) with random s_k^l and s_s, (right) the forecasts $\hat{c}_{frt}$, used in the season-wide simulations, resulting from (12) with random s_k^l and s_s.

The inputs provided to the fruit development model y_B and T_{can}^{24}, originated from the simulations with HPS lighting in Kuijpers et al. (2021), these are depicted in the middle panel of Fig. 5. The value of the fixed time-shift delay in the tsm t_{tsm} was set to 30 *days*, similar to Kuijpers et al. (2021). The value of $\mathbb{D}$ in both the model by Vanthoor and the mVt was set to 8. The models were compared based on their output, the prediction of fruit harvest. Note, however, that the tsm and mVt both predict the fruit harvest due to assimilate partitioning at a particular time instance. To be able to compare the prediction of fruit harvest, the impulse response functions of the latter two models were substituted for Ξ in (3) resulting in a prediction of fruit harvest. Based on the prediction of fruit harvest, the income through yield was calculated. The product price data used in the latter analysis originated either from the average yearly trend μ_{fft} in (10) or the product price data from an arbitrary year between 2015 and 2020. The predictions of the mVt are meaningful when the fruit development dynamics have achieved a steady-state value. The results presented below, for all models, are solely based on the period after the 1st of April.

Table 1 presents the predicted cumulative fruit yield over the simulation interval, expressed in dry matter. One can observe that the predictions by the mVt are close to that of the Vt, a difference of 0.01 $kg.m^{-2}$. The calculation of the income through yield does not only require a good cumulative prediction, but due to the time-varying product price, the prediction itself is also key. Therefore, Table 1 also presents the RMSE of tsm and mVt with respect to Vt, normalised to the RMSE of the tsm. One can observe from the results on the normalised RMSE, that the mVt performs considerably better than the tsm. This is supported by the predicted fruit harvest y_F as shown in Fig. 9. For the sake of clarity, the data for fruit harvest by mVt and the tsm in Fig. 9 were filtered by means of moving average filters with window sizes of 1*day* and 3*days*, respectively. Even though the harvest prediction by the tsm is considerably different from the predictions by Vt, the average value is close to the fruit harvest predictions by Vt. The latter observation is supported by the data in Table 1. One can conclude that the predictions of fruit harvest by the mVt are considerably better than the tsm. Moreover, the predictions of the mVt are close to the predictions by the Vt.

Table 1 – The prediction of cumulative fruit yield over the simulation interval and the RMSE of tsm and mVt with respect to Vt, resulting from the simulations. The RSME presented is normalised with respect to the RMSE of the tsm simulation. The prediction of cumulative yield over the simulation interval is expressed in dry matter. Similar results are presented for the income through yield based on the average yearly trend μ_{fft} and the product price data from an arbitrary year between 2015 and 2020.

	$\int y_F(t)\,dt$ $[kg.m^{-2}]$	RMSE [–]	$\int y_F(t)\mu_{fft}(t)\,dt$ $[€.m^{-2}]$	RMSE [–]	$\int y_F(t)c_{frt}(t)dt$ $[€.m^{-2}]$	RMSE [–]
tsm	2.03	1	18.90	1	19.46	1
mVt	2.15	0.29	19.84	0.73	20.52	0.74
Vt	2.14	0	19.79	0	20.49	0

Figure 10 and Fig. 11 present the income through yield corresponding to the fruit harvest predictions presented in Fig. 9. The predictions of income through yield in Fig. 10 are based on the average yearly trend μ_{fft} and the predictions in Fig. 10 are based on data from an arbitrary year between 2015 and 2020. One can observe that, due to the smooth and slowly varying product price in Fig. 10, the accuracy of the mVt as compared to the Vt increased. The latter observation is supported by the corresponding RMSE in Table 1. The corresponding normalised RMSE for the prediction of income through yield was higher for the mVt as compared to the fruit harvest prediction. As one can observe from Fig. 10, however, the predictions of income through yield by the tsm also improved. Figure 11 and Table 1 also show the predictions of income through yield based on

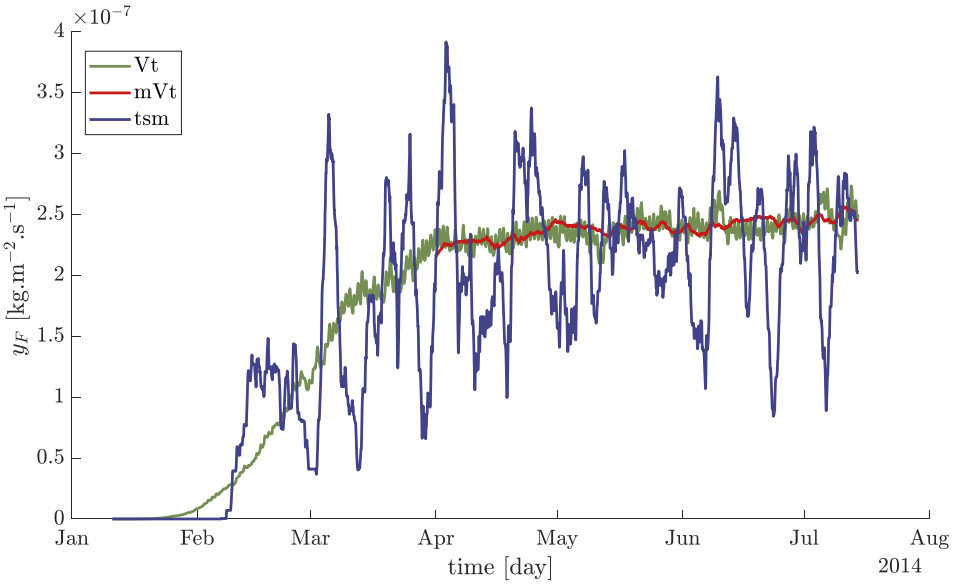

Fig. 9 – Predicted fruit harvest y_F by the Vanthoor model (*Vt*), the modified Vanthoor model (*mVt*) and the time-shift model (*tsm*). For the sake of clarity, the plotted data for *mVt* and *tsm* is filtered by means of moving average filters with window sizes of 1 *day* and 3 *days*, respectively. Only the data for mVt after the 1st of April is shown.

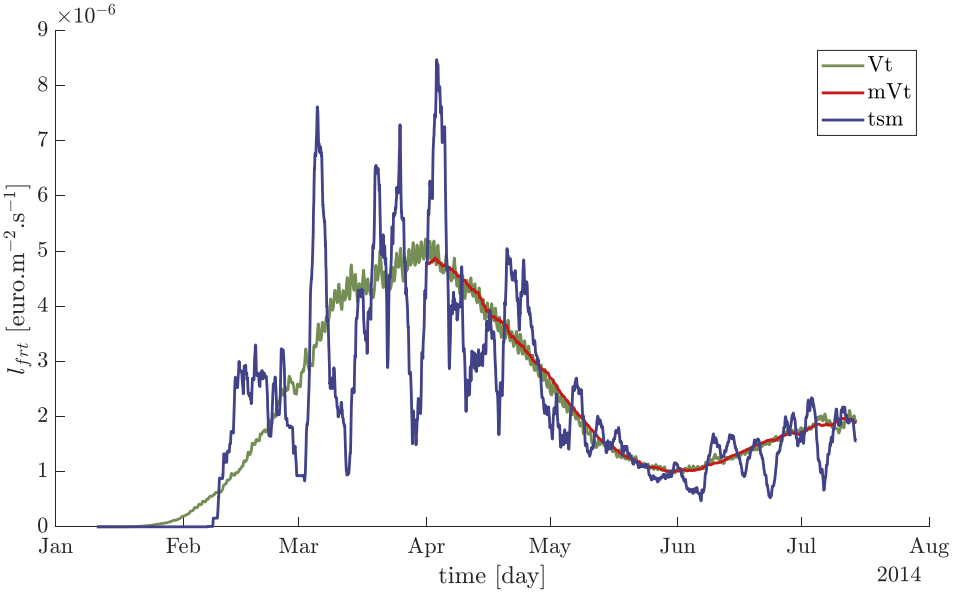

Fig. 10 – Predicted income through yield l_{frt} by the Vanthoor model (*Vt*), the modified Vanthoor model (*mVt*) and the time-shift model (*tsm*). For the sake of clarity, the plotted data for *mVt* and *tsm* is filtered by means of moving average filters with window sizes of 1 *day* and 3 *days*, respectively. Only the data for mVt after the 1st of April is shown. The product price c_{frt} is based on the average yearly trend μ_{fft}.

data from an arbitrary year between 2015 and 2020. One can conclude that the predictions of income through yield of the mVt are close to that of the Vt, hence slowly varying product prices result in better predictions.

3.2. *Uncertainty analysis*

The simulation study presented in this subsection aims to evaluate the effect of the product price forecast error on the control strategy u_d^* which results from (1). The effect of the product price forecast error may depend on the product price, the prevailing weather, and the state of the system. We present the analysis for three time instances, spread throughout the growing season:

- 11th of January. At this day: the prevailing weather represents winter weather, the crop has just entered the productive phase and the product price is high.

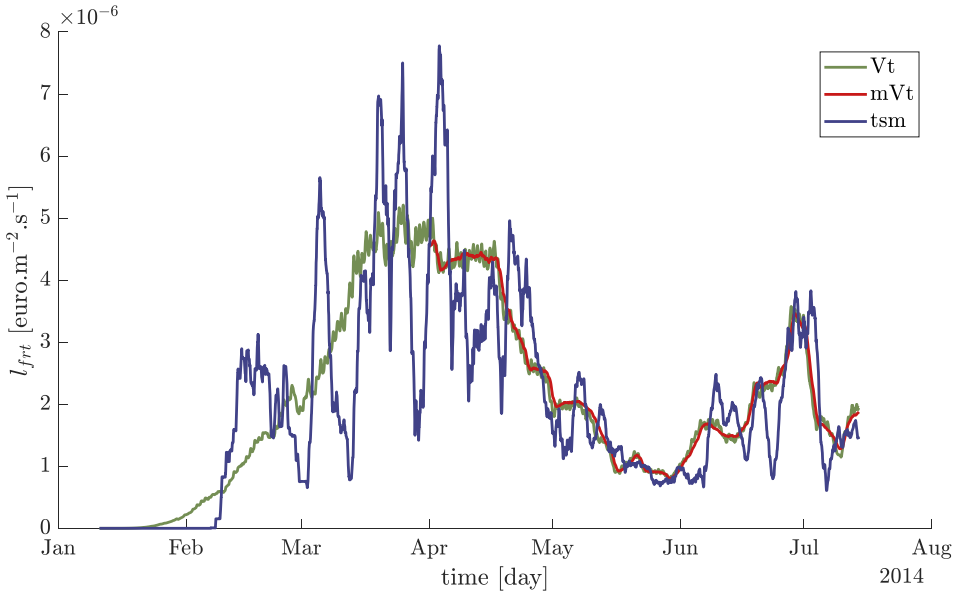

Fig. 11 – Predicted income through yield l_{frt} by the Vanthoor model (*Vt*), the modified Vanthoor model (*mVt*) and the time-shift model (*tsm*). For the sake of clarity, the plotted data for *mVt* and *tsm* is filtered by means of moving average filters with window sizes of 1 *day* and 3 *days*, respectively. Only the data for mVt after the 1st of April is shown. The product price c_{frt} is based on the data of an arbitrary year between 2015 and 2020.

- 10th of June. At this day: the prevailing weather represents summer weather, the crop is well into its productive phase and the product price is close to its lowest point, see Fig. 6, the 161th day of the year.
- 29th of July. At this day: the prevailing weather represents summer weather, the crop is well into its productive phase and the product price increases rapidly in the middle of the FGP, see Fig. 6, the 210th day of the year.

At every time instance, the control strategy was optimised using (1), for 120 different product price forecasts. The mVt model in (9) was employed as fruit development model. The product price forecasts originated from (13) with randomly selected values for the slope, the switching value and randomly selected convergence to the upper or lower bound.

Figures 12, 13 and 14 show the results of the uncertainty analysis for the 11th of January, the 10th of June and the 29th of July, respectively. For every time instance, the set of product price forecasts is presented in the top left graph. Histograms for the accumulated CO_2 injection u_{CO2}, accumulated gas use s and difference in the 24*hr* average temperature on the first day of the prediction horizon are presented in the remaining graphs. These indicators were chosen as these affect fruit development either through temperature and/or CO_2 concentration. The resulting control trajectories do not deviate considerably from those observed in Kuijpers et al. (2021). The product price forecasts in the top left graph were color-coded based on the switching value, s_s in (13), and whether they predict an increasing product price or a decreasing product price. The histogram entries were coloured accordingly. Note that, for the uncertainty analysis at January 11th the optimisation algorithm did not converge for 4 out of the 120 simulations. Figure 12 only shows the 116 simulations in which the solver did converge. An important issue arises when the solver does not provide a feasible solution to the optimisation problem. While we have not experienced this in our simulations, we propose that when this is the case, the control trajectories obtained at the most recent feasible optimisation step are used to provide a control input.

Overall, one can observe that the product price forecast error did not considerably affect the control trajectories of the system. The temperature difference obtained during the first day is significant at all time instances. This result is expected as a different product price forecast was used as compared to the simulations of (Kuijpers et al., 2021) which were used to determine the initial condition. For the temperature difference a clear distinction between increasing and decreasing product price forecasts can be observed at the 11th of January and more clearly at the 10th of June. In general, a lower temperature strategy resulted from the optimisations with rising forecasts. A lower temperature results in a longer FGP. The fruit harvest in these simulations is, thus, prolonged to benefit from the higher prices that were forecasted. The accumulated gas use and CO_2 injection do not show similar differences due to increasing and decreasing product price forecasts. The magnitude of the forecast error, here ± 0.25 *euro*.*kg*$^{-1}$ for the length of one to two months is large compared to practice. The forecasts converge to the 3σ bounds, which represent a confidence bound of $>99\%$. One can conclude from the figures that the change in the control strategy is small compared to the magnitude of the uncertainty simulated here. The temperature difference during the first day might result from of the controller temporarily decreasing the temperature to benefit from the high product price by increasing the FGP. If the temperature, however, is only decreased temporarily, the first assumption in Subsection 2.3 does not hold.

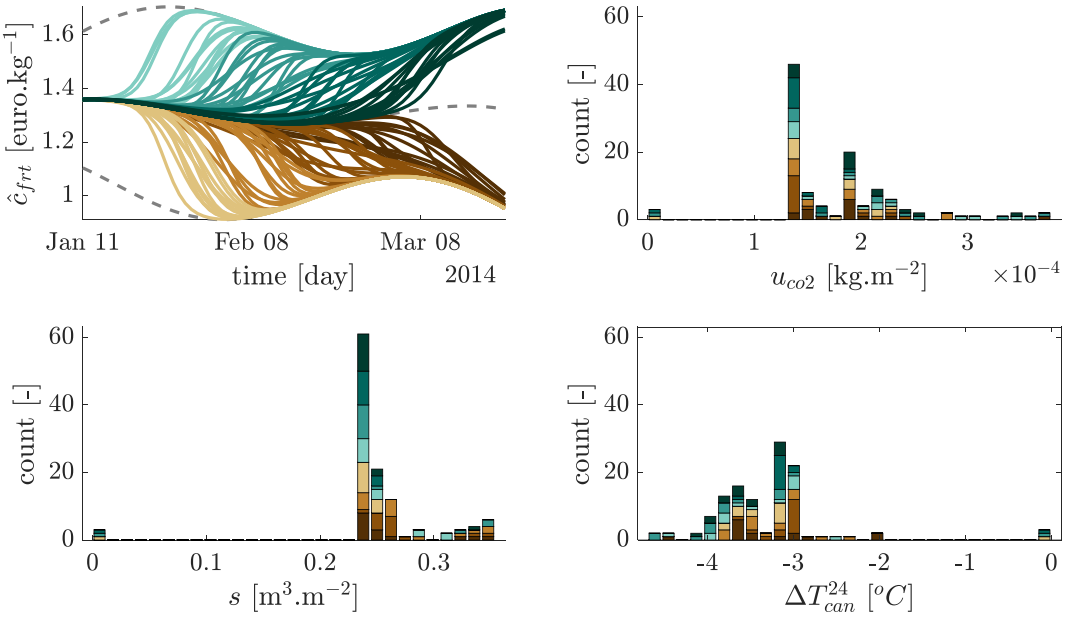

Fig. 12 – Results of the uncertainty analysis at January 11th. Histograms are presented for (top right) the accumulated CO_2 injection u_{CO2} and (bottom left) accumulated gas use s during the first day of the prediction horizon for various price forecasts. A histogram of the difference in the 24 hr average temperature between the start and end of the first day ΔT^{24}_{can} is presented in the bottom right. The entries in the histograms are coloured based on the evolution of the price forecast as presented in the top left graph.

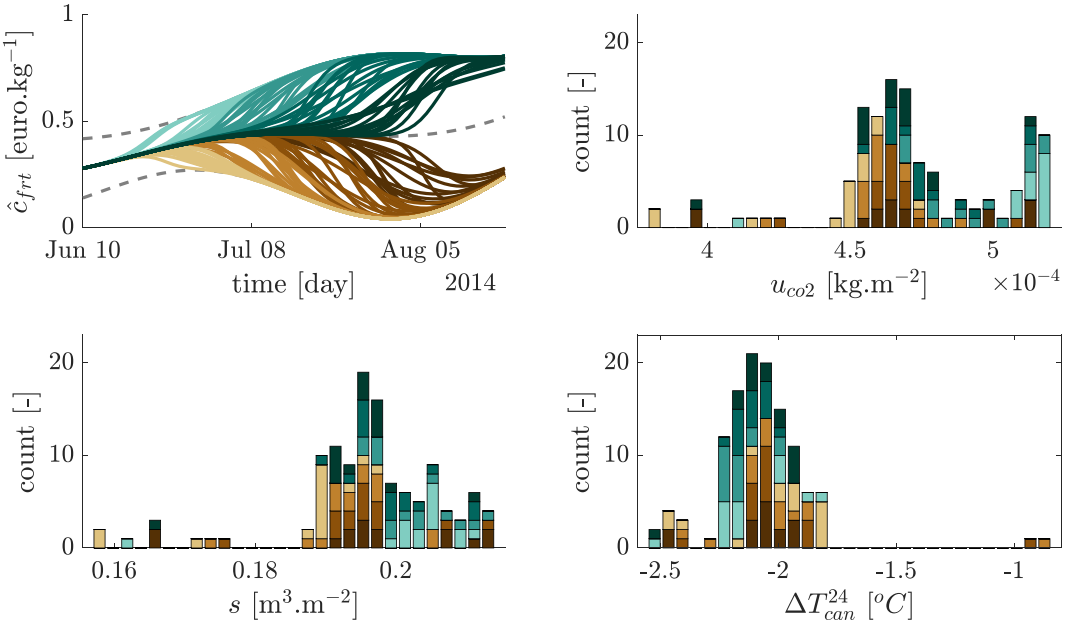

Fig. 13 – Results of the uncertainty analysis at June 10th. Histograms are presented for (top right) the accumulated CO_2 injection u_{CO2} and (bottom left) accumulated gas use s during the first day of the prediction horizon for various price forecasts. A histogram of the difference in the 24 hr average temperature between the start and end of the first day ΔT^{24}_{can} is presented in the bottom right. The entries in the histograms are coloured based on the evolution of the price forecast as presented in the top left graph.

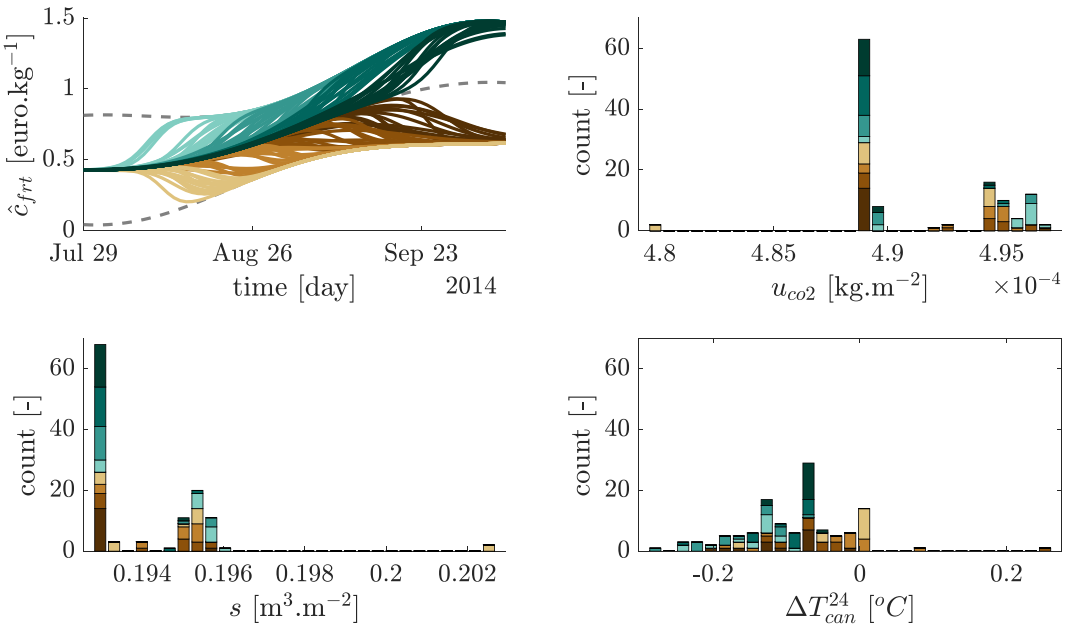

Fig. 14 – Results of the uncertainty analysis at July 29th. Histograms are presented for (top right) the accumulated CO_2 injection u_{CO2} and (bottom left) accumulated gas use s during the first day of the prediction horizon for various price forecasts. A histogram of the difference in the 24 hr average temperature between the start and end of the first day ΔT^{24}_{can} is presented in the bottom right. The entries in the histograms are coloured based on the evolution of the price forecast as presented in the top left graph.

3.3. *Season simulation*

The simulation study presented in this subsection encompasses closed-loop simulations in which the control strategy resulted from (1) with either the tsm or mVt fruit development model. For each of the configurations, two season-wide simulations were run, i.e. one with the product price forecasts according to (12) and one in which the forecast matches the average yearly trend.

The results presented in this subsection are based on data after the 1st of April, i.e. when the fruit development dynamics have attained a steady state. In order to obtain a realistic prediction of fruit harvest and income through yield, the optimised control strategies are applied to the Vanthoor model. The value of $\mathbb{D}$ in both the model by Vanthoor and the mVt was set to 8.

The optimised trajectories for the inputs of the fruit development model from simulations with the product price forecast from (12) are presented in Fig. 15. The price at the current time instant $\bar{c}_{frt}$ is based on data from an arbitrary year within the data by GFactueel (2020). One can observe a higher 24hr -average temperature T^{24}_{can} in the control strategy based on the mVt, i.e. the average temperature over the interval is 22.0 °C compared to 21.7 °C for the control strategy based on the tsm. Also, an increase in the amount of assimilates partitioned to the fruits can be observed, i.e. the total amount of assimilates partitioned to the fruits in the control strategy based on the mVt is 4.70 $kg.m^{-2}$ compared to 4.49 $kg.m^{-2}$.

The inputs of the fruit development model presented in Fig. 15 were input to Vt, the resulting outputs are shown in Fig. 16. From Fig. 16, one can observe a higher fruit harvest and income through yield in the simulations with the mVt as compared to those with the tsm. Similar results were obtained with the simulations based on the average yearly trend. The resulting resource cost l_{res}, income through yield l_{frt} and operational return J that have been obtained during the simulations are presented in Table 2. One can observe that the mVt realised a higher income through yield and a higher resource cost. The resulting operational return for the simulations with mVt was lower than the tsm. The latter was unexpected as the mVt is hypothesised to describe the behaviour of the actual fruit development model (the Vanthoor model) in a better way. Boundary effects at the end of the observed interval might be the cause. Concluding, the differences in operational return are low. The control strategy employed in the simulation with the mVt model is, however, considerably different from the simulation with the tsm model.

Upon comparing the rows in Table 2, conclusions can be drawn on the effect of different product price forecasts. The difference in operational return between the two forecasts was 0.04 €.m^{-2}, less than 1% of the total performance realised in this interval. Depending on the product price at the current time instant, the forecasted product prices are considerably different.

4. Discussion

In Subsection 4.1, the performance of the modified Vanthoor fruit development model is compared to the model by Vanthoor (2011) and the tsm. In Subsection 4.2, limits to the applicability of the proposed approach are discussed. In Subsection 4.3 and Subsection 4.4, the suggested performance increase upon implementation of RHOC, as stated in Kuijpers et al. (2021), is evaluated.

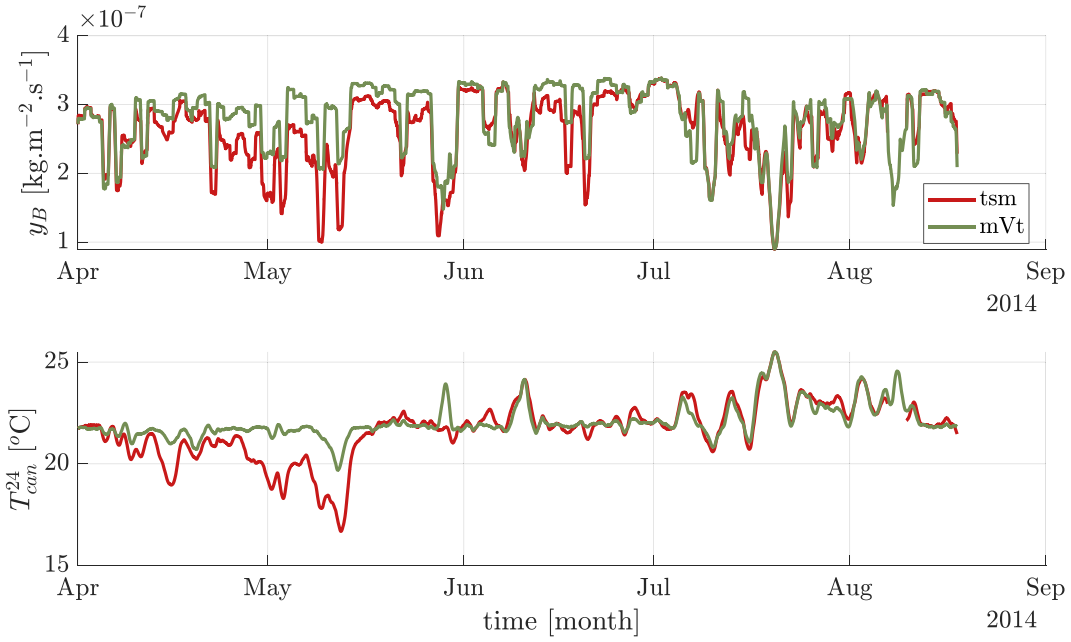

Fig. 15 – Optimised trajectories for (top) assimilates partitioned to the fruits y_B and (bottom) 24 *hr*-average temperature T^{24}_{can} using the time-shift model (*tsm*) and the modified Vanthoor model (*mVt*) in the optimisation algorithm. The product price forecasts $\widehat{c}_{frt}$ originate from (12), the price at the current time instant results from GFactueel (2020).

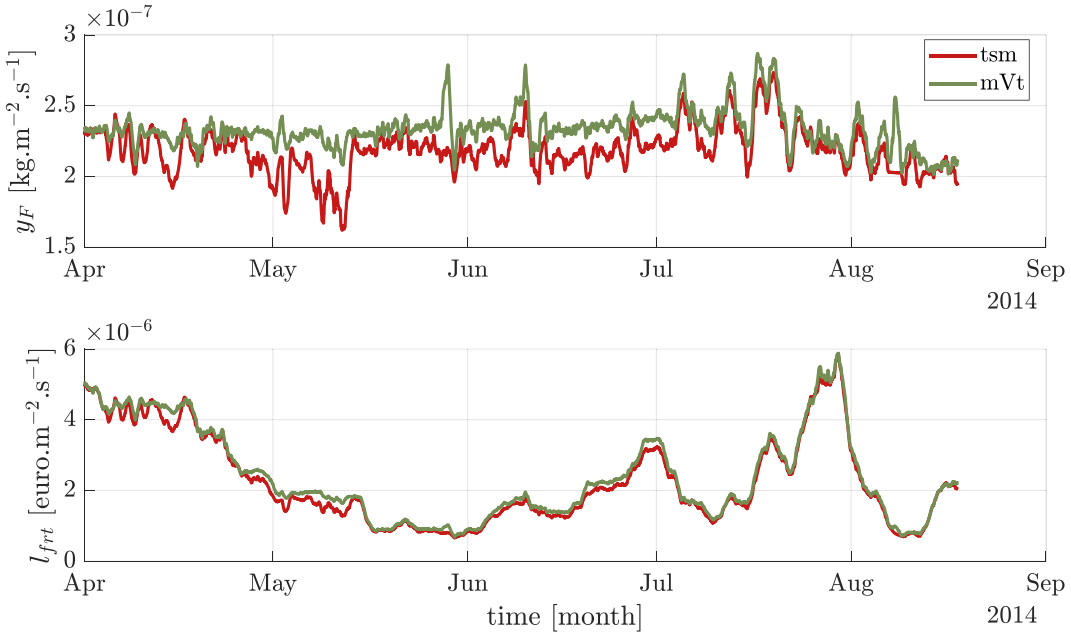

Fig. 16 – Outputs of the original Vanthoor fruit development model using the inputs as presented in Fig. 15 for (top) predicted harvest y_F and (bottom) income through fruit yield l_{frt}. The trajectories result from simulations with the inputs from the simulations with the tsm and mVt. The signals presented here have been filtered by a moving average filter with a window of a day for the sake of clarity.

4.1. *SRHOC with fruit development*

In this subsection the results are discussed focusing on the differences between the tsm model and the mVt model and the mVt model and the model by Vanthoor (2011).

One can observe from Table 1 that the predictions of fruit harvest and income through yield of the mVt were closer, to the original model by Vanthoor (2011) than the tsm. The largest difference can be observed in the prediction of fruit harvest y_F, the RMSE of the mVt with respect to the model by Vanthoor, is

Table 2 – Total values over the interval for the cost of resources l_{res}, the income through yield l_{frt}, the operational return J and carbon footprint p_2. The columns indicate the fruit development model which is used in the closed-loop simulation, i.e. tsm or mVt. The rows indicate the type of product price forecast that is used, i.e. the average yearly trend or product price forecasts $\hat{c}_{frt}$ that originate from (12).

	tsm				*mVt*			
	l_{res} $[€.m^{-2}]$	l_{frt} $[€.m^{-2}]$	J $[€.m^{-2}]$	p_2 $[kg.m^{-2}]$	l_{res} $[€.m^{-2}]$	l_{frt} $[€.m^{-2}]$	J $[€.m^{-2}]$	p_2 $[kg.m^{-2}]$
Forecast	3.16	27.11	23.95	36.65	5.32	28.63	23.31	62.17
Average	2.28	26.18	23.91	26.40	5.28	28.55	23.27	61.90

29 % of the respective RMSE of the tsm. Concluding, the mVt model outperforms the tsm in prediction accuracy with respect to the Vanthoor model. The prediction of fruit harvest due to assimilate partitioning by the mVt is more dispersed throughout the FGP compared to the tsm, as can be observed in Fig. 4. Consequently, the mVt was less sensitive to short fluctuations in the product price forecast.

In Table 1 and Fig. 9, one can observe that the predictions of fruit harvest and income through yield of the mVt were close to the predictions by the model by Vanthoor. This suggests that the effect of the two assumptions presented in Subsection 2.3 was not significant in the simulation. Note, however, that it is not possible to simulate the performance of the system upon integration of the model by Vanthoor into (1). It is, therefore, not possible to compare the mVt model to the model by Vanthoor in closed-loop simulation. By using the mVt instead of the model by Vanthoor, the prediction horizon can be decreased from over 60 *days* to 3 *days*. This 94% reduction in the number of decision variables in the optimisation problem in (1) results in approximately 94% less constraints and a reduction in the Hessian and Lagrangian sizes of 99%. Concluding, the loss in accuracy due to the presented assumptions and modifications is suggested to be minor. The latter takes into account the context of the application in mind, i.e. automatic control of the greenhouse system and what is enabled by the approach, i.e. SRHOC.

4.2. *Applicability*

In this subsection, limits to the applicability of the proposed fruit development model in combination with SRHOC are presented.

Due to the second assumption presented in Subsection 2.3, the dynamics in the fruit development system are assumed to be in a steady state. The proposed fruit development model therefore only provides meaningful predictions if the fruit development dynamics in the real system also achieved this steady state. According to the original fruit development model by Vanthoor (2011), the dynamics in the fruit development system achieve a steady state after approximately two and a half months of generative growth, see Fig. 5. Concluding, the proposed fruit development model provides meaningful predictions after two and half months of generative growth. The latter assumption is added to the assumption of the model presented in Kuijpers et al. (2021), i.e. that the crop has fully developed its leaf area. The latter two assumptions require an experienced and high-tech grower to maintain the actual crop in this steady state through crop management, i.e. leaf and fruit harvest. To what extent deviations from this steady state affect the prediction accuracy should be the topic of further research.

The SRHOC approach presented in Subsection 2.2 is based on the fact that the prediction of fruit harvest y_F is not required in the definition of the optimisation problem in (1). If, however, the calculation of the cost function or the constraints depends on the prediction of fruit harvest, the SRHOC cannot be used. For example, a constraint limiting the harvest y_F per day cannot be included in the proposed approach. Also, if a) the assimilates are buffered for a long period of time, as is the case in for example lettuce, or b) the value of the buffer affects the system, cost function or constraints, the proposed approach cannot be employed. Also, the model assumes that the fruit is sold immediately when the FGP has passed, which is not the case in truss tomatoes, as in the latter, ripe fruits are not always harvested directly.

In case the fruits are sold as part of long-term contracts, the value of a single assimilate can be based on the agreed value. Note, however, that if for a contract a certain amount of product has to be produced, one should include this constraint in the optimisation algorithm in (1). This constraint, based on cumulative production would, however, introduce a pure integrator, hence an extended prediction horizon is needed.

Upon implementation of the proposed fruit development model, attention has to be paid to the optimised trajectories for the greenhouse air temperature. The results presented in the uncertainty analysis, more specifically Fig. 13, potentially indicate the controller purposefully controls to a higher temperature resulting in a decreased FGP. Temporarily increasing temperature, however, conflicts with the assumption in Subsection 2.3. If the optimised trajectories for the greenhouse air temperature temporarily increase, a shorter FGP is obtained during that time interval. This mechanism can be used by the controller to shift fruit harvest to time intervals with more beneficial product prices. This control strategy, however, invalidates the assumption in Subsection 2.3, leading to sub-optimal performance. Note that this does not occur with the tsm model, the latter does not model t_{tsm} as a function of temperature. The effect in the mVt model might be mitigated through increasing the time interval over which the average temperature is calculated. This will make a temporary increase in the average temperature use more resources, making this strategy less rewarding. If, however, the temperature interval extends beyond 3 *days*, the prediction horizon has to be extended accordingly.

The simulation results are based on a controller that has full state information. This might be a valid assumption considering

the states of the greenhouse climate system, which can be measured through climate sensors. In practice, crop state information, however, will have to be obtained through the use of observers combined with crop sensors. The use of full state information enables the evaluation of the effects of uncertainty due to price forecast errors, whereas including the uncertainty resulting from state estimation might hinder such an analysis.

4.3. *Effect of time-shift model on performance*

In Kuijpers et al. (2021), an increase in performance of 10 % was observed when comparing optimal control with respect to the current state-of-the-practice represented by Vermeulen (2016, p. 330). The simulation studies in the latter research employ the tsm as fruit development model.

The open-loop simulations with the tsm presented in Subsection 3.1, show that the predictions of fruit harvest are inaccurate with respect to the mVt and the Vt, see Fig. 9 and the RMSE in Table 1. Because of shape of the impulse response, see Fig. 4, the tsm resembles an unrealistic sensitivity with respect to the product price forecast exactly t_{tsm} into the future as compared to practice which will be closer to the mVt impulse response. The effect of this unrealistic sensitivity was also observed in Kuijpers et al. (2021) upon changes in the product price t_{tsm} into the future. The modified Vanthoor model provides a more realistic sensitivity with respect to future product price forecasts due to its impulse response shape, see Fig. 4.

A comparison between the total operational return during an interval of 4.5 *months* with the tsm and mVt in closed loop did not reveal considerable differences. Unexpectedly, the operational return realised by the controller that employs the tsm model in this period is higher than that realised by the controller that employs the mVt model. The resulting inputs to the fruit development model of both simulations were applied to the Vanthoor model afterwards. This seems to contradict the results obtained as part of the simulation study presented in Subsection 3.1, in which the predictions by the mVt model were closer to the predictions of the Vt model as compared to the tsm model. The latter, however, does not guarantee performance in case the model is integrated into (1). This indicates that the performance of the controlled greenhouse system is not considerably affected by using the tsm.

4.4. *Effect of product price forecast errors*

The simulation studies presented in this research suggest that the effect of the product price forecast error on the performance of the controlled greenhouse system is small. This statement is supported by the results of two simulation studies. The uncertainty analysis presented in Subsection 3.2 shows that the difference in the control strategy is minor compared to the magnitude of the forecast errors used. The magnitude of the forecast error used in the latter simulation study was close to $\pm 0.25\ €.kg^{-1}$ for the length of a month. According to the historical data, $>99\%$ of all product price forecast errors lie between these $\pm 0.25\ €.kg^{-1}$ bounds. We envision that forecasts in practice will provide a better prediction than those used in the uncertainty analysis in Subsection 3.2.

The simulations with the two fruit development models in closed-loop also show that the effect of different product price forecasts is small, i.e. a difference of $0.04\ €.m^{-2}$ was obtained in the simulations of both fruit development models. Note, however, that the product price forecasts employed in the simulations both have a forecast error with respect to the actual evolution of the product price. The product price forecast error in the latter situation has to be calculated with respect to the data from GFactueel (2020), as this data was substituted for $\bar{c}_{frt}$ in (12). In this research, the performance of the controlled greenhouse system could not be simulated using the product price from GFactueel (2020). The control algorithm in (1) cannot handle tabular data. The algorithm in (1) symbolically calculates the Hessian and Lagrangian, this is not possible if a data sequence is input (Betts, 1998). In practice, it will depend on the forecasting service whether the forecasts are supplied in tabular form. The simulation study presented in Subsection 3.3 suggests that the effect of different forecasts does not considerably affect the performance of the controlled greenhouse system. As both price forecasts have a different forecast error, the similar performance of the controlled greenhouse system obtained in the simulations suggests a low sensitivity of the performance with respect to the forecast error.

Both simulation studies, discussed here, suggest the performance decrease due to product price forecast error to be not considerable.

5. Conclusion & recommendations

One of the main contributions of this paper is the development of a fruit development model which can be used in combination with a computationally viable short receding horizon optimal controller. Open-loop simulations with the proposed model show an accuracy in predicting fruit harvest similar to the original model by Vanthoor (2011). The proposed fruit development model, however, can be used with SRHOC, i.e. a prediction horizon of 3 *days* suffices compared to the 60 *days* required upon implementation of the original model. The use of the proposed fruit development model is limited because of the assumption that the fruit development dynamics have achieved a steady state, simulations suggest that this happens approximately 3 *months* after the start of generative growth. The proposed algorithm can be used during the remaining part of the season in practice and to simulate the effect of product price forecasts as presented here.

Using the aforementioned fruit development model and artificial product price forecasts the effect of the product price forecast error on the control actions and the performance of the greenhouse has been evaluated. An uncertainty analysis reveals that the product price forecast error does not significantly affect the optimised control trajectories. Season-wide simulations support the hypothesis that the product price forecast error does not considerably affect the performance of the controlled greenhouse system.

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

Authors W.J.P. Kuijpers, E.J. van Henten and M.J.G. van de Molengraft are part of the research program LED it Be 50% with project number 14217, which is supported by the Netherlands Organisation for Scientific Research, LTO Glaskracht, Signify, Ridder Growing Solutions and B-Mex. The authors would like to thank GFActueel for making their historical product price data available for this research.

REFERENCES

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., & Diehl, M. (2019). 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

Betts, J. T. (1998). Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics, 21*(2), 193–207. https://doi.org/10.2514/2.4231

van Beveren, P. J. M., Bontsema, J., van Straten, G., & van Henten, E. J. (2015). Optimal control of greenhouse climate using minimal energy and grower defined bounds. *Applied Energy, 159*, 509–519. https://doi.org/10.1016/j.apenergy.2015.09.012

Bunte, F. H. J., Bolhuis, J., Bont, C. J. A. M., Jukema, G. D., & Kuiper, E. (2009). *Prijsvorming van voedingsproducten*. Landbouw Economisch Instituut. https://edepot.wur.nl/172116.

GFactueel. (2020). *GFactueel marktprijzen—Tomaat tros grof*. Misset Uitgeverij B.V. https://www.gfactueel.nl/Prijzen/.

van Henten, E. J. (1994). *Greenhouse climate management: An optimal control approach*. Ph.D Thesis. Wageningen University. https://doi.org/10.1016/S0308-521X(94)90280-1.

HSL. (2019). *HSL, a collection of Fortran codes for large-scale scientific computation*. See http://www.hsl.rl.ac.uk.

Kempkes, F. L. K., Janse, J., & Hemming, S. (2014). Greenhouse concept with high insulating double glass with coatings and new climate control strategies; from design to results from tomato experiments. *Acta Horticulturae, 1037*, 83–92. https://doi.org/10.17660/ActaHortic.2014.1037.6

Koning, A. N. M. (1994). *Development and dry matter distribution in glasshouse tomato: A quantitative approach* [Ph.D Thesis]. Wageningen University.

Kuijpers, W. J. P., Katzin, D., van Mourik, S., Antunes, D. J., Hemming, S., & van de Molengraft, M. J. G. (2021). Lighting systems and strategies compared in an optimally controlled greenhouse. *Biosystems Engineering, 202*, 195–216. https://doi.org/10.1016/j.biosystemseng.2020.12.006

Pavlov, A., Hunnekens, B. G. B., Wouw, N.v. d., & Nijmeijer, H. (2013). Steady-state performance optimization for nonlinear control systems of Lur'e type. *Automatica, 49*(7), 2087–2097. https://doi.org/10.1016/j.automatica.2013.04.017

Rabbinge, R., & Ward, S. A. (Eds.). (1989). *Simulation and systems management in crop protection*. *Pudoc*.

Ramírez-Arias, A., Rodríguez, F., Guzmán, J. L., & Berenguel, M. (2012). Multiobjective hierarchical control architecture for greenhouse crop growth. *Automatica, 48*(3), 490–498. https://doi.org/10.1016/j.automatica.2012.01.002

Seginer, I., & Ioslovich, I. (1998). Seasonal optimization of the greenhouse environment for a simple two-stage crop growth model. *Journal of Agricultural Engineering Research, 70*(2), 145–155. https://doi.org/10.1006/jaer.1997.0261

Seginer, I., van Beveren, P. J. M., & van Straten, G. (2018). Day-to-night heat storage in greenhouses: 3 Co-generation of heat and electricity (CHP). *Biosystems Engineering, 172*, 1–18. https://doi.org/10.1016/j.biosystemseng.2018.05.006

van Straten, G., & van Henten, E. J. (2010). Optimal greenhouse cultivation control: Survey and perspectives. *IFAC Proceedings Volumes, 43*(26), 18–33. https://doi.org/10.3182/20101206-3-JP-3009.00004

Tap, F. (2000). *Economics-based optimal control of greenhouse tomato crop production* [publisher not identified].

Vanthoor, B. H. E. (2011). *A model-based greenhouse design method*. Ph.D Thesis. Wageningen University.

Velden, N. van der, & Smit, P. (2019). *Energiemonitor van de Nederlandse glastuinbouw 2018*. https://doi.org/10.18174/505786

Vermeulen, P. C. M. (2016). *Kwantitatieve informatie voor de Glastuinbouw 2016-2017*. Business Unit Glastuinbouw: Wageningen University & Research.

Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming, 106*(1), 25–57. https://doi.org/10.1007/s10107-004-0559-y

Xu, D., Du, S., & van Willigenburg, G. (2018). Adaptive two time-scale receding horizon optimal control for greenhouse lettuce cultivation. *Computers and Electronics in Agriculture, 146*, 93–103. https://doi.org/10.1016/j.compag.2018.02.001